

Embryology needs rules, not new laws

The British committee on human embryological innovation cannot accept the professional view that innovation should be free. The best hope is some form of regulation falling short of legislation.

Now that the British Government's formal inquiry into the potential of human embryology has finished taking written evidence (see page 739), the question in the front of the Warnock committee's mind is certain to be that of deciding what should be done to regulate research and, ultimately, its clinical application. As always, there is a huge spectrum of administrative and legal devices from which to make a choice, which must (as always) be constrained by the knowledge that the government of the day will not accept a recommendation it considers to be politically imprudent. At one end of the spectrum is the administratively simplest choice — doing nothing — which governments will reject for fear of seeming complacent. At the other is the politically safest choice — formal advance registration of all intended experiments, consideration of the relevant costs and benefits by some statutory procedure and a licensing system for all experiments and innovations in this field, or perhaps even a piece of legislation declaring that nothing new should happen.

What, in the circumstances, should the Warnock committee recommend? first, there is a simplifying assumption that it should embrace the distinction between laboratory research and clinical practice. This is the foundation for the intelligent argument of the Royal Society's recommendation to Warnock, which boils down to a plea that once it has been decided that a particular embryo will not be implanted into a living uterus, it should be left to ethical committees in laboratories and hospitals to decide what experiments should be allowed, and for how long after fertilization human embryos should be kept alive. Intellectually, that position cannot easily be shaken. The potential benefits of a better understanding of the causes of human infertility and of genetic defect, not to mention an understanding of the course of embryonic development, are so great that responsibly unfettered research should be wholeheartedly welcome. In passing, it may meanly be noted, those who either entertained or inconvenienced the residents of south London suburbs at the weekend by marching on a group of Medical Research Council laboratories at Carshalton, ritually offering a few among their number for arrest, should be the first to applaud the wish of embryologists and geneticists to work with human rather than with other-primate material.

Autonomy

The trouble is that even where only research laboratories are concerned, the "leave-it-to-the-ethical committee" option will not wash politically. The Royal Society's committee giving evidence to Warnock has implicitly based its argument on the proposition that an early embryo is not a living thing, in the sense of being potentially autonomously self-replicating, but rather a kind of passing parasite, dependent on a uterus for full development. The assumption is literally and also logically correct. Moreover, the Royal Society has usefully injected into the debate on what are called "test-tube babies" the observation that human coitus, like that of some but not all other primates, is not confined by endocrine mechanisms to the times of female ovulation, with the consequence that germ cells (male or female) may accumulate genetic defects as they age, in advance of fertilization. Plainly the mechanisms by which these errors are naturally corrected cry out for investigation and for the most severely practical reasons — to improve, so to speak, on nature.

None of this will weigh with those who do not accept the starting premise, who hold that all early human embryos are living things and who insist that decisions about their fate can be made only by people willing to be accused of murder. That is the wellspring from which the ferocity of the anti-abortion lobby springs. Moreover, the existence of this awkward if misguided point of view cannot be ignored. Certainly politicians must be conscious that constituents holding this contrary point of view (most but not all of whom will be Roman Catholics) will not accept the Royal Society's opinion or the consonant opinions of others without a great deal more argument. Indeed, many of those agreeing with the Royal Society committee's view that a human being is an animal that has somehow survived for forty weeks in a human uterus would not cherish the notion that such a substantial body of dissenters, wrong-headed though they may be, should be forced by legislation or the lack of it into compliance. Sooner or later, the scientific community will have to take up this argument, taking on the Vatican if necessary in the process. Meanwhile, is it too much to ask that the Warnock committee should give the argument a hearing?

Seemliness

There is therefore a case for an interim arrangement on experimental investigations of human embryos. The immediate need is that there should be some means of making sure that investigations are carried out in a seemly fashion, not in such a way that they give gratuitous offence. The Royal Society's proposal, sensible though it is, that local ethical committees should decide between sensible and other experiments, will not in the circumstances suffice, technically because ethical committees do not exist in all interested laboratories, more substantially because there is no way of telling from the outside whether an ethical committee is simply a facade behind which a principal investigator and his associates may hide and also because, properly in a free society, there is no law against the conduct of research by people who do not belong to the research community as defined by the Royal Society's *Yearbook*. So, as a temporary expedient, the Warnock committee should propose some system for the registration of experiments with local safety committees before they can be undertaken, and some means by which information about the totality of the experiments carried out in any year can be made publicly known. But would not that be simply an exercise in public relations? In retrospect, has not that been the sole function of the British Genetic Manipulation Advisory Group, soon (it is to be hoped) to go out of business?

Yes is the answer, but so what? If such a stratagem were to make research with human embryos socially palatable when, otherwise, it would be a focus for heated controversy and perhaps illegal, would not the consequences be beneficial for all concerned, researchers and those likely to benefit from what remains to be found out? It is true that the ultimate beneficiaries of a better understanding of human embryology (most of whom, by definition, are not yet alive but whose putative patents and grandparents stand proxy) should know better. But that is no more relevant than the assertion that unemployment would cease to be a politically important problem if only monetarist economics were more widely understood. In human embryology as in monetarist economics, the onus lies with the protagonists to argue the case



that makes their practices acceptable.

The Warnock committee should therefore pass part of the buck with which it has been stuck back to the academic embryologists. What is to be made of the applications of the new techniques? The obvious difficulty for Warnock is that in Britain medicine differs from that in most other industrialized states in that those who practise it do not have to be registered with the state. Homoeopathy flourishes, osteopathy is in good shape and there is no law to prevent one person telling another that influenza may be cured by an infusion of herbs picked from some hedgerow at full moon. To be sure, people are regularly sent to gaol for passing themselves off as registered medical practitioners, or for extorting money from those from whose desperation in illness they unfairly win financial gain, but there is nothing to prevent private people from operating sperm banks, AID clinics or the like except the general injunction against advertising, the chance that they would be less likely than qualified physicians to recruit patients and the financial risks of being sued for damages in an unsympathetic court if something went wrong.

Warnock cannot follow the advice offered by the Royal College of Obstetricians and Gynaecologists that all establishments at which *in vitro* fertilization is carried out should be registered under some new legislation without tilting at a long-established principle underlying British medicine. But how, then, could it consistently draw a line between *in vitro* fertilization or embryo transfer on the one hand and the now long-established practice of AID? Here again, the best course will be to ask that all embryonic manipulations should be discussed with some ethical committee having a statutory existence, that appropriate data should be recorded confidentially but in such a way that they could afterwards be recovered and that there should be some means of letting the world know, after the event, what has been done. The simple way of doing that is to set up a formal group along the lines of the genetic manipulation committee to exert a general supervisory influence on local ethical committees, to consider difficult questions as they arise and to keep the general public informed of what is going on. There could be no assurance, of course, that the need for such a committee would melt away. In genetic manipulation, the need for the committee disappeared as evidence accumulated that the dangers of research had been pitched too high. In embryology, ordinary people's beliefs can be changed only by public argument. □

Which pipers to pay?

The United States should find a better way of paying the true cost of research.

THE time at which the grant-making organizations in the United States should have come to a decision about indirect costs has long since passed. The issue is the degree to which research grants by agencies such as the National Institutes of Health (NIH) should be inflated by an amount that will compensate the recipient university or institution for the time and trouble involved in spending the money. That universities incur extra costs when their academics succeed in the increasingly cut-throat competition that passes in Washington for a market in bright ideas is not disputed, in which sense the agencies' admission that indirect costs must somehow be met is but an acknowledgement that academic research needs two kinds of support, the direct cost of research projects and some element of overhead cost. So why (see page 742) is there now in prospect another round of argument about the percentages that NIH will henceforth pay, on top of the nominal value of research grants, to recipient institutions? Two obstacles stand out — the diversity of recipients (state but also private universities) and of donors (NIH, the Pentagon and so on). Most European governments meet the need differently, by subsidizing possible recipients with too little discrimination. The ideal probably lies in the proposal canvassed but then defeated in Congress in the 1960s that there should be general subsidies to research institutions with a proven record of success. Why not dust it off? □

Peace makes problems

Britain's domestic argument about nuclear weapons is getting nasty, with both sides in the wrong.

WITH a British general election perhaps only a few weeks and at the most a few months away, it is understandable but also unforgivable that the Secretary of State for Defence, Mr Michael Heseltine, should have attacked the British Campaign for Nuclear Disarmament (CND) on the grounds that many of its leading members manifest "the calculating professionalism of full-time socialists and communists". For the organization has its roots in the heady days of the late 1950s, when the late Mr Hugh Gaitskell persuaded the then-putative Labour government (eventually elected in 1964) to stick by the collaborative defence of Western Europe by means of nuclear weapons if necessary. That argument was so fiercely but openly conducted, and so educative even for those not directly involved, that the stripling CND was regarded not as a subversive organisation but as a kind of necessary, but exceedingly polite, consequence of a resounding defeat. Its fortunes quickly faded.

Since then, the campaign has made its way in the world by marching from one defence establishment to another, usually in the rain at Easter. Its membership has been revived since the decision at the end of 1979 by the Council of the North Atlantic Treaty Organization that it would be prudent from the end of 1983 to base intermediate-range nuclear weapons in Western Europe if an agreement with the Warsaw Pact states had not by then been reached. The campaign is now loudly protesting that Mr Heseltine's charge of "left-wing domination" is a "smear", which is in the strict sense true. In states such as those of Western Europe, in which there is no law against communist, let alone socialist, opinions, to complain that left-wingers dominate elected committees of voluntary organizations is strictly speaking beside the point. The same complaint might just as well be made against several trade unions in Western Europe.

The recent general election in West Germany is a vivid pointer to the issues with which organizations such as the campaign are concerned. From the particular to the general, they are as follows:

- Given accepted assumptions about strategic policy, do the practical arrangements foreseen make sense? In Britain, many people (including at least one former Foreign Secretary, Dr David Owen) answer in the negative to the proposition that Trident should replace Polaris submarines, for example.

- Whatever the arrangements for coordinated defence, does not a national government's possession of or hospitality for nuclear weapons diminish the chance of immunity from attack? (The simple answer is yes, the more complicated is the question "What if everybody concluded that?")

- Are not nuclear weapons such abominations that they should be foresworn? (Of course: but arms control hangs fire.)

Western European opinion against nuclear weapons is untidily organized around one or more of these questions. Governments would be better advised to find better ways of answering them than to broaden their attack in constitutionally uncomfortable ways.

The peace, or protest, movements also need to be more circumspect. On 19 April, an offshoot of the campaign called the "CND Sizewell Working Group" published a pamphlet called *The Plutonium Connection*, intended as a basis of its evidence against a proposal to build a nuclear power station in Suffolk. Mercifully, on the basis of a decision by the Sizewell inspector last week, there is now a good chance that the case will be inadmissible, but how can an organization supposedly dedicated to the admittedly urgent question of strategic arms spare so much effort to argue against a nuclear reactor that cannot function within the decade? And how can an organization with purportedly clear motives join in, as CND will be doing this week in London, a "Green Rally" whose objective seems to be to form a party of Greens on West German lines? Single-cause pressure groups merely invite smears when they, like their attackers, stray from the central issues. □

US nuclear power

Court gives backing to state moratorium

Washington

THE Supreme Court dealt the beleaguered nuclear power industry a sharp blow last week by ruling that state governments possess extensive powers to regulate or prevent the construction of new power plants, provided only that they do not trespass on the federal government's exclusive responsibility for radiological safety.

In a unanimous opinion, the court upheld the legality of California's 1976 statute declaring a moratorium on future reactor construction. The federal government had joined two Californian utilities in challenging the statute on the grounds that the 1954 Atomic Energy Act vested the regulation of nuclear matters solely in the federal government.

Justice Byron White, accepting California's contention that it had declared the moratorium for economic and not safety reasons, said the Atomic Energy Act made the federal government responsible for all nuclear safety questions but left states with their "traditional responsibility" for regulating electric utilities on the basis of need, reliability, cost and other local concerns. He added that the act does not expressly require the states to construct or authorize nuclear power plants or prohibit the states from deciding not to permit the construction of any further reactors.

The court's opinion is particularly ominous for the nuclear power industry because of antinuclear sentiment among state legislatures. Five states have already passed laws similar to California's and 30 states supported California in its arguments before the Supreme Court. A *Washington Post*-American Broadcasting Company News poll this month found 65 per cent of a random sample of Americans opposed to building new nuclear power plants and only 27 per cent in favour.

Justice White's opinion also dashed the hopes of some in the industry that it would be possible to overturn state laws barring new plant construction by proving that the statutes had in fact been motivated by fears on safety and not on economic or other grounds. The court said it ought not to become embroiled in an unsatisfactory attempt to discover the "true motive" of state legislators, and pointed out that safety and cost questions were closely linked.

If not properly stored, the court declared, nuclear wastes might leak and endanger the environment and human health. "The lack of a long-term disposal option increases the risk that the insufficiency of interim storage space for spent fuel will lead to reactor shutdowns, rendering nuclear energy an unpredictable and

uneconomical adventure. The California laws at issue here are responses to these concerns."

The Californian law does not ban the construction of new power plants permanently, but only until the state is satisfied that an adequate national system exists for the disposal of high level nuclear waste. The Department of Energy has begun the search for a permanent geological depository for spent fuel but is not expected to select a site until 1987. And the site might not become operational until the end of the century.

Ironically, the ruling came while the nuclear industry was celebrating another Supreme Court decision a day earlier absolving the Nuclear Regulatory Commission (NRC) of any need to investigate the psychological impact that local residents might suffer following a decision to restart the undamaged Unit 1 reactor at Three Mile Island.

Residents of Harrisburg, Pennsylvania, had been supported by a federal appeals court in maintaining that under the Environmental Policy Act of 1969, NRC was obliged to include psychological fallout in its mandatory assessment of the impact on the environment of a decision to restart. But that ruling was struck down unanimously by the Supreme Court, which said the relationship between a construction project and the anxiety it might cause was too indirect to be part of a federal review procedure.

Justice William Rehnquist said risk was a pervasive element of modern life and many risks were generated by modern technology. But the fears associated with these risks were for Congress to consider in policy decisions; the 1969 law was not intended to give citizens a general opportunity to air their policy objections to proposed federal actions.

The ruling means that NRC can proceed to a decision on restarting Unit 1 without conducting extensive enquiries about the psychological aftermath of the Three Mile Island accident. NRC officials are still reviewing technical problems at the plant — particularly the condition of the steam generators — and will not say when they will decide on a restart date.

In his Supreme Court opinion, Justice Rehnquist made a point of saying the court did not intend to denigrate the fears of Harrisburg residents or deny their existence. In fact, since the Three Mile Island accident, a number of studies, including several commissioned by NRC, have suggested that psychological stress and the use of tranquilizers and alcohol have increased substantially.

Peter David

Spanish oil deaths

Toxin is elusive

Barcelona

PHYSICIANS and scientists seem to be agreed that the cause of the toxic oil syndrome (TOS) that killed 300 people in Spain in 1981 is some constituent of unpurified rape-seed oil. But the identity of the toxic material is still obscure.

TOS was the subject of two meetings last month in Madrid — the meeting of the CSIC (Spanish Consejo Superior de Investigaciones Científicas) Programme for the Study of TOS and the World Health Organization Working Group on TOS. The weeks preceding these meetings were marked by sit-ins and demonstrations by many of the 20,000 people affected by the epidemic, some of them with severe chronic disabilities, asking the Spanish Government for more medical and social assistance.

The epidemiological evidence about the origin of TOS was reviewed by the WHO Working Group which found "overwhelmingly" in favour of the association of TOS with consumption of illegally sold denatured rape-seed oil. Other hypotheses had been formulated as the origin of this new and complex disease, such as a purely infectious agent or a link with tomatoes treated with certain pesticides or with toxins produced by fungi associated with grape-seed oil, but the evidence has been unconvincing. This international endorsement of denatured rape-seed oil as the culprit is important for Spanish scientists working on the subject, coming as it does after accusations by those affected, the media and government officials that they had not considered other alternatives.

Another important question discussed was the toxic agent responsible for TOS. The CSIC programme concentrated its research on fatty acid anilides, which were present in up to 2,000 p.p.m. in the most reliable samples of toxic oil available for analysis. The evidence for the toxicity of anilides was reviewed last year (*Nature*, 298, 608; 1982) and new correlations between anilides and lesions in model animals similar to those observed in patients were presented. The WHO working group recommended that the presence of anilides in at least 700 mg l⁻¹ should be taken as one of the markers for toxic oil but the evidence for the anilides as the origin of TOS was considered still inconclusive.

The need for increased and coordinated research to understand the origin and consequences of TOS was one of the main conclusions of the two meetings. The medical recommendations stress the need for Spain to develop a national epidemiological programme, to prepare for publication all the epidemiological evidence and to define a long-term follow-up programme for the affected people.

Pedro Puigdomenech

Data falsification trial

Drug testing lab was "shambles"

Washington

FOUR scientists who held key positions at the major testing laboratory Industrial Bio-Test went on trial in Chicago last week on charges of falsifying safety data on drugs and other products submitted for government approval. In the government's opening statement, Assistant US Attorney William Spence painted a picture of a laboratory "in shambles", with mice escaping, watering equipment malfunctioning and flooding laboratories, large numbers of animals dying and left to rot in their cages, and too few trained technicians to carry out necessary tests. Spence charged that to hide its failings and to meet deadlines, the laboratory had fabricated results, cut tests short, and then, to cover its tracks, back-dated reports, destroyed records, and lied to investigators from the Food and Drug Administration (FDA).

The indictment names four specific products, all since retested and approved. FDA has begun reviewing 800 other tests carried out by the laboratory, however, and the Environmental Protection Agency is reviewing 200 tests performed on pesticides. According to the government, Industrial Bio-Test was one of the largest independent testing laboratories in the United States, until FDA began its investigation in 1976. At that time, it was carrying out as many as 2,000 tests a year.

The government's case will rely heavily on the testimony of employees who have been granted immunity from prosecution. One key witness will be Philip Smith, a technical writer whose job it was to write up reports on the studies carried out at the laboratory. In Smith's earlier testimony to a grand jury, he told how data from a test of the drug Naprosyn, an anti-inflammatory agent developed by Syntex Corporation, could not be found but nonetheless turned up — handwritten — in a final draft of the report. Smith and the laboratory's chief of clinical pathology had already concluded from a search of records that the work in question — blood and urine analyses — had never been done. Smith left that section of the report blank and told Paul Wright, director of rat toxicology and a defendant in the case, that he would not sign the report; several weeks later, Smith saw the final draft and saw not only the handwritten insertions but also found that a signature page from an unrelated study had been appended, making it appear that he had signed the report.

Smith also told the grand jury that he had been ordered by another defendant, Moreno Keplinger, previously the laboratory's manager of toxicology, to manipulate data reported on studies of two agricultural chemicals, Sencor and Nemacur, developed by Chemagro. According to Smith, Keplinger was dissatisfied with the data reported on the positive con-

trols (animals fed a known carcinogen in order to demonstrate the susceptibility of the test strain to cancer) saying they did not show a high enough incidence of cancer. Keplinger is alleged then to have ordered Smith to incorporate into the final report a control group from an unrelated test, even though the controls had not been fed the carcinogen, as required by the Sencor and Nemacur studies, but had had it painted on the animals' skin. The government also claims the studies ran for only 14 months, not the 18 months required, and that records were doctored to conceal this.

The fourth substance mentioned in the indictment is TCC, an antibacterial agent used in deodorant soaps and manufactured by Monsanto. According to the government, Wright, who left Industrial Bio-Test to join Monsanto, brought pressure on the laboratory to change a pathology report which found harmful effects of TCC on the testicles of exposed rats. The government claims that the other unfavourable data were suppressed as well.

Spence, in his opening statement, dwelt at length on slipshod conditions at the laboratory. He described a chaotic scene of animals dying or escaping and being replaced in the middle of tests. Spence claimed that as many as 20 or 30 mice would escape from their cages each day (he said they were in rat cages and could slip through the bars) and would be periodically rounded up in "chloroform hunts" by employees armed with squeeze bottles of chloroform. The captured mice were returned to cages almost at random, he said, since there was no marking or numbering system.

As well as Wright and Keplinger, the company's president, Joseph Calandra, and the senior group leader for rat toxicology, James Plank, are also accused in the case.

The defence plans to question the government's use of immunity for its witnesses and what defence attorneys call the government's "bullying" and "brainwashing" to obtain evidence. Judge John Nordberg has already denied a pre-trial motion to dismiss the charges for this reason, but the defence will certainly focus on this in the trial. The defence attorneys have also accused the government of violating attorney-client privilege by offering immunity from prosecution to two attorneys who earlier represented the laboratory during the FDA investigation. The laboratory itself, however, is not charged in the case. Nordberg's denial of a motion to dismiss on this count is under appeal.

The defence will also argue that what the government calls "horror stories" about the laboratory's conditions in fact reflect acceptable laboratory practice at that time. The trial is expected to last three to four months.

Stephen Budiansky

Laboratory safety

Bending rules to aid research

ARE laboratory chemists more at risk than other people because many of the chemicals they handle are toxic, or does their competence give them immunity? This was the essence of the question discussed two weeks ago in London at a meeting organized by the Royal Society of Chemistry for the Federation of European Chemical Societies (FECS).

Insurance statistics from West Germany show that the accident and occupational health record of chemists compares very favourably with other research and industrial professions. And the evident longevity of members of the Royal Society of Chemistry was put forward as supporting evidence.

A view widely held was that the competence of the laboratory chemist in handling dangerous chemicals, and the fact that very small quantities are generally involved, should be taken into account when legislation on health and safety is formulated. Most of the regulations on the use of dangerous chemicals in European countries make no distinction between large-scale use of chemicals in industry and small-scale use in the laboratory. According to Professor M. Corn (School of Hygiene and Public Health, Johns Hopkins University, Baltimore), the US Occupational Health and Safety Administration is contemplating relaxing some of its more stringent requirements for specialized research laboratories.

Much of the discussion at the London meeting, however, centred on harmonizing legislation on an international scale. Recommendations included the standardization of statistics on the reporting of accidents and near misses, a research programme to establish what precisely constitutes a carcinogen, safety audits to be extended to universities (Dr N. H. Pearce of the University of Bristol reported a blackly humorous case of conflicting legislation which led the university first to install a door with two-hour fire resistance, and then, by adding mandatory view-panels and louvres, to reduce its efficacy to a few minutes), architectural research aimed at the better planning of laboratories (Dr Istvan Szentpeteri of the Chemical Engineering Centre, Budapest, pointed out the hazards of the high-rise laboratory blocks, popular among Hungarian planners for prestige reasons), and the incorporation of compulsory courses on toxicology and safety procedures into undergraduate studies in chemistry.

The media were subjects of one resolution, which called for less emotive reporting when discussing alleged carcinogens or other toxic chemicals.

Vera Rich

Human embryo experiments

Societies urge a softer line

A LIBERAL set of rules to govern *in vitro* fertilization of human embryos is being advocated by the Royal Society and the Royal Society of Obstetricians and Gynaecologists (RCOG). In evidence to the Inquiry on Human Fertilization and Embryology under Mrs Mary Warnock, set up by the British Government last year, both organizations say the implantation stage is too soon in human development to be taken as an end-point for research.

Both professional societies would thus go further than the provisional guidelines published last year by the Medical Research Council (MRC). But RCOG also argues that premises at which *in vitro* fertilization is carried out for clinical purposes should be licensed by the government. Among other things, the Warnock committee has been asked to say whether legislation will be necessary (see *Nature* 299, 475; 1982).

The Royal Society's evidence, produced by an *ad hoc* group under Professor R.L. Gardner, goes further than MRC by considering the use of *in vitro* fertilization in research screening for genetic abnormalities and in basic research in human embryology. Like MRC, the group insists that informed consent should be obtained from donors of eggs and sperm and that embryos produced for experimental purposes should "under no circumstances" be implanted for development *in vivo*. The Royal Society differs from MRC however, in saying that "it would be restrictive to have to specify the precise research purposes for which any given specimen of human embryonic material is obtained".

The report says that decisions about how long *in vitro* development is allowed to proceed should be left to local ethical committees. It advocates "a degree of flexibility" in the regulation of such work, with the value of information to be obtained from a particular experiment as a primary consideration, as well as how long normal embryonic development should be allowed to continue. The research possibilities discussed, and approved in principle, include the investigation of teratogenic effects of drugs and viruses, studies of metabolism and gene expression in early development and the genetic manipulation of embryos.

The RCOG report, produced by its Ethics Committee under the chairmanship of Dr J.D. Loudon, approves the practice of *in vitro* fertilization and of embryo replacement, or the use of ova from a donor for subsequent implantation in a host mother. But it says that neither technique should be used except for a stably cohabiting heterosexual couple.

RCOG suggests that a statutory body be established to advise the government on related questions, to suggest who be empowered to inspect premises and to keep a register of directors of institutions where

the techniques are employed. The committee approves research into freezing of embryos for storage, with a time limit, but recommends that embryos should be kept only for a foreseeable future purpose. It says that human embryos produced for research should not be maintained beyond the early neural development stage (day 17 after conception).

One of the more difficult ethical conundrums raised in this evidence is whether the new techniques should be offered to all couples who request this assistance and who give fully informed consent, even when there are known risks. RCOG concludes that such decisions should be left to the individual physician, but the question has been highlighted by a recent exchange of letters in the *British Medical Journal*. Dr Patrick Steptoe and Mr Robert Edwards, who pioneered *in vitro* fertilization in Britain, severely criticized Dr Alan Trounson and others in Melbourne for using eggs from a 42-year-old infertile woman for im-

plantation in a host mother. The fetus spontaneously aborted and was found to have a chromosomal abnormality.

Edwards and Steptoe, well used to criticism themselves, wrote of their suspicions of "hurried decisions taken under pressure" and argued for an age limit on donors for embryo replacement because of the greater risks of abnormalities with increasing age. They felt that the host mother's own eggs should have been used. Trounson *et al.* replied that all parties had given fully informed consent to the procedures and that his group respected the patients' right to make decisions. Dr Edwards last week reiterated his view that fully informed consent should not be a sufficient condition for offering embryo replacement.

The Warnock Committee has now completed taking formal written evidence, although it is not expected to report before next year. The British Medical Association has another committee on the question which is reporting shortly on ethics in the field of *in vitro* fertilization but its recommendations are unlikely to be as liberal as those offered by the Royal Society and RCOG.

Tim Beardsley

Cancer research

Support for charity sustained

PEOPLE'S wish to see cancer cured is stronger than ever — that seems to be the message of the 1982 report of the Imperial Cancer Research Fund (ICRF), one of the top three medical charities in Britain. ICRF's income last year was £18.8 million, 10 per cent more than in the previous year and, in real terms, larger than at any time in the past decade.

Spending on research, largely at ICRF's London laboratories in Lincoln's Inn Fields, decreased slightly in real terms but remained at the comparatively high level of 1981. Even so, ICRF managed to increase its accumulated capital from £31.4 to £34.5 million.

Two new projects and the fruition of another are likely to alter the balance sheet in future years. The laboratories being built at South Mimms, just north of London, will cost "many hundreds of thousands of pounds each year" once they start operating under the direction of Dr Thomas Lindahl, probably in 1985.

Lindahl will move there from the ICRF laboratories in Mill Hill, London, which are due to close down soon. Other senior staff will move to a new developmental biology unit at the University of Oxford's department of zoology with Professor Richard Gardner as honorary director. In return for use of the laboratories, ICRF is contributing building costs.

The other new project is a laboratory at St Bartholomew's Hospital in the city of London, devoted to the pursuit of oncogenes. The new laboratory, costing £500,000 to construct and £200,000 to run each year, will open this autumn with Dr

John Wyke of ICRF as director; he will also become director of the medical oncology unit at St Bartholomew's which is already maintained by the fund.

Although Wyke's move will double the size of his group and leave him well placed to initiate studies on oncogenes in human tumours — an area conspicuous by its absence from the ICRF programme — he has no immediate plans to do so. Instead, he and his collaborators plan to stick to much of their present programme of research, concentrating on animal models in an attempt to discover the normal functions of the genes that are occasionally converted into oncogenes.

Launching the annual report last week, Dr Walter Bodmer, scientific director of ICRF, drew particular attention to a new association of a hormone abnormality with an increased risk of cancer that is emerging from a long-term prospective study of the female population on Guernsey.

The study, which has involved the analysis of blood and urine samples of 13,000 women over 22 years, has revealed a "highly significant deficit" in the excretion of androgen metabolites in the 30 women who subsequently died of ovarian cancer. Although the conclusions are still tentative, Dr Jack Cuzick, the statistician who has made the study, estimates that the increase in relative risk of ovarian cancer in women with low androgen excretion is two- to three-fold, about the same as that already established for a group of Guernsey women with hormonal abnormalities and at increased risk of breast cancer.

Peter Newmark

UK universities

New blood posts make bad blood

THE British Government has now announced the final allocation of the 312 new appointments to be made in universities this year under its "New Blood" and Information Technology schemes (see *Nature* 301, 7; 1983). Details of the schemes, which will cost £100 million over three years, were first given last December by the Secretary of State for Education and Science, Sir Keith Joseph, and required that universities should apply to the University Grants Committee (UGC) by the end of February for an allocation from the national pool.

Several universities had to abandon their normal procedures to meet the deadline. In some cases, academic or faculty committees have not been fully consulted, as they would ordinarily expect to be. The allocations were processed in record time by UGC and the research councils. But UGC (which is getting used to criticism) has now been accused of bad management and illogicality by some universities.

Seventy of the new posts were originally set aside for lectureships and research posts in information technology. The remainder, for "new blood", comprises 32 arts and social science research posts and 210 for other sciences, among which the physical sciences and engineering predominate. More controversially, the lion's share of

chancellor of the University of Salford (two appointments) said UGC had chosen not to notice how different universities had responded to the budget reductions since 1981. Salford has been hailed as an example of how universities may survive through collaboration with local industrial companies. Professor Ashworth said the allocations were reminiscent of 1981, when some newer technologically-based universities were particularly hard hit.

UGC said last week that the committees making the allocations took into consideration the procedures used by different universities in ranking their applications. Although several of the universities were

clearly hoping that the new appointments would restore ground lost since 1981, UGC said that the effects of the 1981 cuts had not been taken into account except in the practical sense that departments must have adequate resources to support a new post.

A UGC member said that while some of the choices may have been less than perfect, UGC had learned a lot, and that this would be reflected when next year's allocations are made. It was also pointed out that, although many good applications had to be turned down, this was only the first year of a three-year scheme. But there had been no systematic bias, and some of the aggrieved institutions should be least consider the possibility that their applications "simply weren't good enough". They should, it was said, "try harder next time".

Tim Beardsley

Agricultural research

UK pressures towards autonomy

BRITISH agricultural research seems to be heading for a shake-up. That is the obvious reading of what the new chairman of the Agricultural Research Council (ARC), Lord Selborne, has been saying since his appointment as acting chairman last August. British farmers would be well advised to contribute towards the cost of agricultural research, Lord Selborne suggests, so as to insulate the enterprise from the vagaries of government policy.

Among the five publicly supported research councils in Britain, the agricultural council has been consistently the least favoured by successive governments. In the late 1960s, official discontent with the research programme of ARC provoked the Rothschild inquiry into the management of British civil science. Lord Rothschild recommended that nearly a third of ARC's budget should be transferred to the Ministry of Agriculture on a commission basis.

Now, as then, ARC is in a cleft stick; either its research makes such a direct contribution to the prosperity of agriculture that the beneficiaries — British farmers — should pay the cost, or its objectives are so distant as to be subsumed within a much more academically-oriented research council such as the Science and Engineering Research Council.

Consistently, ARC has asserted in successive annual reports that its interests lie with agricultural productivity. It may, however, have become the victim of its own propaganda, dependent on what farmers are prepared to pay rather than on taxpayers in general.

Small steps towards the devolution of ARC have been evident for several years. ARC institutes threatened because of financial stringency with closure have usually sought survival by means of farmers' subventions. As a consequence, part of the cost of the programme of research at the Long Ashton research sta-

tion in Wiltshire (south-west England) is paid for by means of a levy on the output of cider and perry manufacturers. Attempts to support research such as that under way at East Malling in Kent (south-east England) on the backs of apple and pear growers have, however, been less successful.

The most recent assault on ARC's survival was that of the Advisory Board on the Research Councils, made public (without precedent) towards the end of 1982, which recommended that the ARC budget should slowly shrink from the financial year beginning in April 1985. The ARC secretary, Dr Ralph Riley, "dissociated" himself from that proposal, and has now published a defence of ARC's position in the council's newsletter *ARC News* (Spring 1983).

Dr Riley says that if the intended cuts are implemented, 500 jobs in council establishments will have to go. *Pro rata*, a third of the posts concerned would be those of scientists. ARC is also, however, hoping to square its budget by encouraging more support for its research from farmers, to which end Riley promises central support for those in outlying establishments who wish to approach financially compliant (or simply rich) research partners.

In these circumstances, ARC may well find it convenient to cast its net more widely than in the past, encompassing not merely the science of how best to grow crops but that of how crops should best be consumed, as food.

To tell from how members of the council talk, partial independence along these lines is now the best solution; the question begged is simply that of whether governments such as the British can institutionalize the subsidy of research, as in the constitution of ARC, without implying in the process that if it succeeds, it should be paid for by somebody else — and that if it fails, it must have been a mistake that it was ever begun. □

UNIVERSITY OF OXFORD

'New Blood' appointments in Science (including Clinical Medicine and Mathematics)

Applications are invited for the following university lectureships tenable from 1 October 1983. It is intended that these appointments should be held in conjunction with a college fellowship. Further particulars of the university lectureships, and of relevant college posts, may be obtained from the head of the department indicated in each

New blood — old institution

the new appointments has gone to the old-established universities of London, Cambridge, Oxford and Edinburgh. The science posts are each worth £20,000 annually to the universities that receive them (out of which has to come a salary for a young researcher). Some of the smaller universities feel that more of the allocations should have gone to departments hit by cutbacks over the past two years.

There were 2,250 applications for the 342 posts, and UGC says that its selection criteria included the age structure in the departments concerned and the likely contributions to scholarship of the proposed post. Universities were also asked to attach a judgement of priority to their applications for science posts, but now several universities have expressed surprise about the subject allocations actually approved. Thus the University of Stirling, which has received three appointments, described the approvals as "having the appearance of randomness". Others were more forthright. Professor John Ashworth, vice-

Polish academics

More sackings come to light

ALTHOUGH martial law in Poland was suspended at the end of last year, several cases are still pending in the courts for offences committed during this period. For the most part, under martial law, the heads of institutes or laboratories tried to minimize the hardship suffered by academics and scholars arrested for continuing trade union activities by granting them unpaid leave of absence for the duration of their captivity. (Those "merely" interned as a preventive measure at the start of martial law were officially assured of job reinstatement.)

Recently, however, Western biologists have been made aware of the case of a young Polish colleague, Jerzy Bal, who has been dismissed from his post at the Warsaw Medical Academy while still in pre-trial detention. The evidence against Bal apparently consists only of denunciation by informers, so it seems unlikely that he will receive more than a suspended sentence. In that light his dismissal before his case has even been heard seems unnecessarily severe.

Unfortunately, Jerzy Bal had the bad luck to be arrested in connection with what is rapidly developing into a show trial. He was arrested last December as a suspected member of the Interfactory Workers' Solidarity Committee (MRKS), the most radical of the grass-roots underground Solidarity groups set up in the early days of martial law. More than 200 people were arrested on suspicion, of whom 21 were detained pending investigation. Nine of them have so far been brought to trial, but proceedings were halted in March, after a showdown in which the defence counsel, Andrzejewski, was dismissed for contempt for calling the proceedings a "kangaroo court". Comments in the army daily newspaper *Zolnierz Woinosci* make it clear that the proceedings are designed as an indirect trial of Zbigniew Bujak, the head of the five-man underground Solidarity leadership, and are closely linked with the forthcoming trial of the five "intellectual advisers" of Solidarity from the dissident workers' Defence Committee (KOR) which formally handed over its civil rights commitments to Solidarity in September 1981.

Under these circumstances, the medical academy may well have been panicked into wishing to dissociate itself from Bal, although his friends in Warsaw stress that during police searches of his flat, no evidence was found to link him with MRKS. It would be unfortunate, however, if the pusillanimous attitude of the medical academy were to set a precedent for other academic institutions whose staff fall foul of the security authorities.

Vera Rich

Solar System exploration

NASA spells out plans

Washington

THE Solar System exploration committee of the National Aeronautics and Space Administration (NASA) has published its recommendations for a "core programme" of planetary missions to be flown before the end of the century. New features of the programme include using derivatives of commercial Earth-orbital spacecraft for inner planet missions ("Planetary Observers") and a new, modular spacecraft named Mariner Mark II which can be modified for a variety of missions to the outer planets, comets and asteroids. The first four proposed missions are:

- **Venus Radar Mapper.** To be launched in 1988, and included in the Reagan Administration's 1984 budget proposal, this mission will provide a radar map of the planet's surface with a resolution of approximately one kilometre. The spacecraft will also carry a radar altimeter to provide topographic information and a gravity sensor.

- **Mars Geoscience/Climatology Orbiter.** This mission, for launch in 1990, would be the first to use the Planetary Observer spacecraft. Its purpose is to orbit Mars for one martian year gathering information on the planet's surface composition, magnetic fields (if any) and seasonal cycles of carbon dioxide, water and dust that interact between the surface and atmosphere.

- **Comet Rendezvous/Asteroid Flyby.** The new spacecraft, Mariner Mark II, is recommended to rendezvous with a short period comet as it approaches the Sun and

then recedes into deep space. Mariner would carry cameras and remote sensing instruments to study the gas and dust boiling off the comet's nucleus, and sensors to measure the gas and carbon dust composition. Three suitable targets have already been identified — comets Enke, Tempel 2 and Honda-Mrkos-Pajdasakova (HMP) — which could be encountered in the mid-1990s after first flying by a main-belt asteroid.

- **Titan Probe/Radar Mapper.** Proposed for launch on some date between 1988 and 1992, this mission would examine the orange-coloured atmosphere that shrouds Saturn's largest moon, and its surface, which is cold enough for methane pools or possibly oceans. The Titan probe could parachute instruments through the atmosphere and send back radar images of the surface. The mission, if upgraded, would also utilize a full-scale Saturn orbiter to study the planet's atmosphere, magnetosphere, satellites and rings. NASA says the combined probe and orbiter would be an ideal candidate for international collaboration.

Excluded from the core programme on the grounds of cost, but still under study by the committee, are missions such as the return of samples from Mars, rovers exploring the martian surface and a buoyant station floating in Titan's atmosphere.

The committee plans to publish the full details of the augmented planetary exploration programme at the end of the year.

Peter David

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

A three-dimensional computer-generated model of the 15-km high Arsia Mons volcano on Mars based on data from Viking Orbiter. Photo: NASA, April 1983.

NIH research grants

Battle on indirect costs continues

Washington

A PROPOSAL by the National Institutes of Health (NIH) to slice \$72 million next year from the amount they pay to universities as reimbursement for the "indirect costs" generated by federally sponsored research has touched off an acrimonious argument between NIH and the major research universities.

Indignant complaints about the proposal dominated discussion last week at a private meeting of the Association of American Universities (AAU), which represents the 50 leading universities in the United States. One participant, Stanford's president Donald Kennedy, called the proposal "a unilateral abrogation of enduring understandings" which was likely to destroy the hitherto close relations between the universities, NIH and the biomedical investigators who receive NIH funds.

The federal government and the universities have been squabbling for years about the mechanism for paying the indirect costs — such as rent, heat and administration — which universities incur as a result of carrying out federally funded research. But the latest proposal from NIH, which has been forwarded to the appropriations committees of the House of Representatives and the Senate to form part of their 1984 budget deliberations, appears to have shifted the debate into a higher gear.

In a report to the committees, NIH complain that indirect costs have been consuming an ever-greater fraction of the total funds available for grants; in 1982, for example, indirect costs accounted for 30 per cent of the funds awarded for research project grants by the Public Health Service compared with about 20 per cent in 1972.

The report describes this rate of increase as "disproportionate" and proposes a number of ways in which indirect costs could be contained. One method, apparently favoured by the institutes, is to freeze the percentage of research money that universities could claim for indirect cost reimbursement at a level equivalent to the average level applied to each institution over the past decade.

Both the proposal and the manner in which it has been brought to Congress have

infuriated AAU. Donald Kennedy told fellow presidents at last week's meeting that NIH had postponed consultation with the universities as long as possible and then issued an invitation to a meeting at three days notice, on the eve of NIH's budget reauthorization hearings in Congress. Especially because those able to attend the meeting had failed to persuade NIH to change their mind, Kennedy said that it is "misleading — indeed disingenuous" for NIH to suggest that their proposals stem from consultation with the universities.

More worrying, from AAU's point of view, is NIH's belief that the rise in the indirect cost fraction of their university research spending is somehow disproportionate. The universities argue that they spend real dollars on indirect research costs and are reimbursed only after going through exhaustive federal accounting procedures. There are, AAU claimed in a recent letter to NIH, "perfectly good explanations" for the increases in indirect cost rates.

One reason, the letter says, is that the nature of biomedical research has changed, with more equipment sharing, larger and more complex research teams and therefore more pooled costs attributable to indirect costs. In addition, direct costs have not kept pace with the cost of conducting research while many items in the indirect cost category, such as energy, have risen faster than the general rate of inflation. Meanwhile, universities have become more efficient in identifying indirect costs attributable to research.

According to AAU, the NIH scheme for containing rising indirect costs would unfairly penalize some universities. Dr Kennedy argues that some state universities claim lower indirect cost reimbursements because they cannot keep the money themselves but send it on to the state government. Such universities have therefore had little incentive to justify realistic overhead rates. The NIH policy, therefore, would be hardest on those universities (mainly private) that had kept careful track of indirect costs and done the accounting necessary to allocate them.

The universities intend to ensure that NIH's proposal is stopped in Congress, as was a proposal last year to impose a 10 per cent across-the-board cut. But for NIH, the arguments for pressing ahead with their attack on indirect costs are compelling. The institutes are already having to pare many 1984 programmes to find the money Congress wants so that the number of investigator-initiated research grants can be kept at 5,000, not reduced to 3,700 as proposed by the President in January. Indirect costs offer a relatively painless way of squeezing out a portion of the extra \$140 million needed for this purpose.

Peter David

US accelerators

Time to choose a giant

Baltimore

EVEN as particle physicists were last week taking up positions for the impending skirmish over the choice of the next large accelerator project at the American Physical Society meeting here, they were being dressed down by White House science adviser George Keyworth. Using unusually sharp language, Keyworth complained that "during the years American physicists have squandered on a pork-barrel squabble, the Europeans have moved boldly ahead".

There are now four large accelerator projects under consideration, and a meeting of the High Energy Physics Advisory Panel (HEPAP) has been arranged for June to recommend one of them to the Department of Energy. Keyworth urged last week that the time had come to choose "the best tools", not to dissipate resources by parcelling them out in response to "popular demand" and that the physics community should commit itself to the best proposal "and stick with it".

In a thinly veiled reference to Brookhaven's foundering ISABELLE project (now renamed CBA), on which the Department of Energy has already spent

The "desertron"

Chicago

THE "desertron", proposed last summer by Leon Lederman and Robert Wilson, would use relatively weak 2.5 tesla superconducting magnets in an enormous 150-km-circumference ring. At 3.5 tesla, the magnets could use stable and easily-machined iron cores to shape the field, rather than the coils themselves as in the more powerful 4.5 tesla Fermilab magnets.

Wilson argued at a meeting last summer that the lower cost of the magnets would more than offset the cost of the extra real estate, especially if the accelerator were built in the desert. The simpler magnets would also require less maintenance, permitting them to be buried permanently in narrow pipes where they could be maintained and aligned by robots.

At a meeting attended by 40 accelerator physicists last month at Cornell University, some of Wilson's assumptions were questioned. The group looked into two other designs for obtaining a 20-TeV-per-beam collider, one using 5 tesla magnets and the other 8 tesla, and concluded that they could be competitive economically with Wilson's original "desertron". Both would use smaller rings than the desertron.

"The cost uncertainties are large enough that all three designs are in the running", said James Bjorken of Fermilab. "These studies are still very preliminary."

Larry Arbeiter

Correspondent's coup

DR Robert Walgate, European Correspondent of *Nature* since 1980 and one of the journal's staff for the past five years, has been awarded one of this year's three Glaxo fellowships for science writing awarded by the British Association of Science Writers. The award, which is worth £1,000, is for Walgate's 20-page "Science in France" contribution to *Nature* published in May last year. □

Candidates for the US "accelerator of the 90s"

Accelerator	Institute	Collisions	Energy	Comments
CESR II	Cornell	e ⁺ /e ⁻	100 GeV	Would compete with LEP at higher luminosity
CBA (ISABELLE)	Brookhaven	p/p	800 GeV	To achieve very high luminosity for probe of rare events
Dedicated collider	Fermilab	p/p, e/p	4 TeV	To extend energy domain before the end of the decade to initiate e/p collisions
"Desertron"	?	p/p, p/p	20 TeV	Vastly to extend energy domain

\$150 million, Keyworth said that "this has to be a time for statesmanship, not for pet projects". According to many particle physicists, however, the Department of Energy may be locked into ISABELLE by the need to save face even though the accelerator would be obsolete before it could be completed.

The four projects among which HEPAP will choose were described last week by Leon Lederman, director of Fermilab (see table). He said that the Cornell proposal (CESR II) would be a duplication of LEP, the electron-positron collider being built at CERN in Europe. The machine is not regarded as a major contender, while Lederman also noted that one option for HEPAP is "none of the above".

According to Lederman, there is nevertheless a strong case for going ahead quickly with at least one of these higher-energy machines. He said this was now accepted not only by experimentalists like himself, and one of his slides bore the legend **IN THE PAST FIVE YEARS, THE OVERCONFIDENCE OF THEORISTS HAS GIVEN WAY TO PUBLIC CON-**

FESSIONS* OF A NEED FOR EXPERIMENTAL GUIDANCE. The accompanying footnote read **EXCEPT GLASHOW.**

But now even Glashow, Lederman said, has tacitly admitted the need for more experimentation in his somewhat whimsical "oilatron" proposal (see *Nature* 3 March, p.6). Lederman said that Glashow's "cockamamie scheme" to build an accelerator to search for oil deposits was "just a disguise — he's not interested in finding oil, he just wants the data".

The proposal that Keyworth is said to favour — and his remarks last week backed up this view — is the "desertron", first discussed last summer. It would use "low-tech" magnets and an enormous ring which, presumably, could only be built in the desert in the south-western United States. This machine, with its energy of 20 TeV per beam, would put the US capability head and shoulders above the "European peril", as Lederman called it, even after CERN's "LEP-prime" (a proton-antiproton collider at 6–10 TeV per beam) is completed, perhaps in 1993.

Stephen Budiansky

Trust counts bird casualties

THE British Trust for Ornithology (BTO) this week celebrates 50 years of monitoring bird populations in the British Isles. The trust's surveys, which are widely respected, are based on information collated from reports sent in by volunteers all over Britain. That they have gained scientific respectability makes a pleasing illustration of the very useful role that amateurs can still play in serious research.

At the trust's headquarters in Tring, Hertfordshire, a small professional staff funded largely by the Nature Conservancy Council coordinates the two main threads of BTO's work. These are the Common Birds Census, which monitors population levels on the basis of amateurs' sightings, and the National Ringing Scheme, which gathers information on mortality and migration patterns. In addition the trust publishes occasional detailed surveys on particular species, and has undertaken studies for the EEC and for BP Petroleum Development Ltd on the effects of oil spills on seabirds. Prompt information from BTO on the effects on bird

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

W.S. Paton/RSPB

BTO figures showed that massive numbers of razorbills, puffins and guillemots were killed off the east coast of Britain in storms during February. populations of the severe winter of 1981–82 was influential in bringing about a temporary ban on the shooting of wildfowl while populations recovered. The trust gains significant revenues from sales of its various publications, and is now seeking sponsorship for a Winter Atlas of distribution of birds in Britain.

Tim Beardsley

US nuclear physics

Argonne's loss

Meanwhile, Argonne National Laboratory lost the first round in its bid to build a new 4 GeV electron accelerator last week when a panel of the Nuclear Science Advisory Committee (NSAC) endorsed a competing proposal submitted by a consortium of 20 south-eastern universities (see *Nature* 17 March, p.204). Although a final decision need not be made until next January, when the budget for 1984–85 is drawn up, NSAC's recommendations have usually carried weight.

The panel's recommendation came as an especially severe blow to Argonne which had been hanging its hopes on the electron accelerator in its effort to regain lost prestige in basic nuclear physics. Argonne director Walter Massey appeared before NSAC last Friday to offer a polite but vigorous dissent from the panel's report.

Massey argued that on purely technical grounds of cost effectiveness, support facilities and staff expertise, Argonne had the edge on the Southeastern Universities Research Association (SURA). And he took exception to the emphasis the panel appeared to place on SURA's pledge to create 35 new tenured faculty positions in nuclear physics if its proposal were approved. Massey noted that this was a policy matter, not a technical matter, and further that "if this is to be a determining factor, then I am convinced that we, in the Midwest, can also respond" with university faculty positions. He said soundings of the University of Chicago and the University of Illinois were favourable.

Massey asked the committee to forward the report to the Department of Energy (DoE) without an endorsement, so that "the institutional and policy questions can be discussed and debated within the proper framework, which has to be at the Department level". Unstated was the likelihood that Argonne may have more influence at that level than SURA.

The SURA proposal calls for an electron accelerator with an energy range of 0.5–4 GeV, to be built in Newport News, Virginia. The panel noted the flexibility that SURA's design offers in being able to expand to 6 GeV (while admitting that there is as yet no justification for going to these higher energies) and SURA's more "conservative" design that some panel members felt was less susceptible to stability problems than Argonne's.

But the panel at the same time expressed some reservations about the SURA proposal, most notably the remoteness of the proposed site from a major university and from an international airport.

NSAC appears ready to forward the panel's report, but may raise some questions in a letter to accompany it. The final outcome is not likely to be settled until DoE makes its budget recommendations, months from now.

Stephen Budiansky

Radiation doses

SIR — The National Radiological Protection Board may well have been courageous in projecting 13 fatal cases of thyroid cancer from the Windscale accident (*Nature* 17 March, p.207) but the accuracy of this result is subject to serious doubt. By relying on the concept of the man-sievert, the need to "look up the dose-response relationship" is eliminated although a straight line is implied which denotes proportionality between dose and incidence. The admission that even in such a large population the radiogenic cancers are not detectable, together with the absence of accepted scientific justification for the "linear hypothesis" of radiation carcinogenesis (especially when extrapolated to minute doses) weakens this exercise more than the other limitations mentioned in your report.

Until radiobiology advances well beyond its current capabilities we won't know what it was that we did not detect.

HARALD H. ROSSI

College of Physicians and Surgeons
of Columbia University,
New York, USA

SIR — I was surprised to read in your issue of 17 March (p.207), "Resurrecting a nuclear accident", that you find the unit man-sievert "chauvinistic". Units cannot all be named after British scientists. I am sure you would not condone an insult to the Swedish nation, so perhaps it is the apparent maleness that your reviewer cannot stomach. Would (s)he rewrite the poets: the child is parent to the person? The National Bitch and Dog Owners Association will be the required title before we know where we are and the Horse-and-Mare-Race Betting Levy Board.

D.K. BEWLEY

London W11, UK

Koestlers' choices

SIR — In your leading article "The worm in the bud" (*Nature* 10 March, p.93) you equate suicide with the assertion of personal freedom. You go on to suggest that through his suicide, Arthur Koestler was directly responsible for his wife's premature death. These premises lead you to conclude that he should not have committed suicide. In my opinion, your judgment fails to address the essential issues in the case.

First, in saying that Arthur Koestler exercised his personal freedom by ending his life, you assume that Cynthia Koestler would have preferred her husband to live on. The short time that Arthur Koestler had left to live would have brought his wife the agony of seeing his increasing debilitation by disease. She may have agreed with him on the desirability of his suicide.

Second, we cannot assume that Cynthia Koestler's suicide was, as you claim, a "consequence" of Arthur Koestler's

suicide. According to all reports, she had an extraordinarily strong attachment to her husband. She might have chosen to commit suicide even had he died naturally.

Finally, no one but Cynthia Koestler herself could decide whether to end her life. To assign to Koestler responsibility for his wife's decision to die with him diminishes her dignity as a human being.

No one can know the reasons for any suicide. If we must make moral judgment at all, we can at least try to appreciate the complexities of each case. Instead, your article simply reaffirms accepted attitudes of disapproval towards suicide. As an epitaph to Koestler your article could hardly do better than to project the conventional viewpoint. It seems fitting that he should die as he had lived: at the leading edge of evolving social values. M.R. VAN SCHRAVENDIJK
Bethesda, Maryland, USA

SIR — I deplore your editorial remarks on the suicide of Mr and Mrs Koestler (*Nature* 10 March, p.93). If matter and energy are interchangeable then death has no meaning other than a change of focus or of phase. Any sense of deprivation we feel is purely selfish. In my view, and those of other members of EXIT (now the Voluntary Euthanasia Society) Mr Koestler and his wife are at liberty to decide when and how they move out of the world in which we can see them: that is "life". Society, and those remaining, are neither "enhanced" (as you put it) nor degraded by their action.

ANITA AIROLDI

London N16, UK

First use of *biologie*

SIR — A modern American textbook¹ on the history of biology (or life sciences) says "The term 'biology' was first introduced at the very beginning of the nineteenth century and publicized by the writings of the French zoologist, Jean Baptiste Lamarck (1744–1829), and a German naturalist, Gottfried Treviranus (1776–1837)." In fact Gottfried Reinhold Treviranus used the word in his *Biologie oder Philosophie der lebenden Natur für Naturforscher und Ärzte* (six volumes, Göttingen 1802–1822²).

Lamarck first used the term *biologie* in his work *Hydrogéologie, ou recherches sur l'influence qu'ont les eaux sur la surface du globe terrestre*, published in Paris in December 1801 or January 1802^{3,4}. Lamarck in his book divides the "physics of the Earth into three essential parts, the first being a theory of the atmosphere, or *Meteorology*, the second, a theory of Earth's crust or *hydrogeology*, and the third, a theory of living organisms, or *biology*"⁴. The planned book on *biologie* never appeared but was replaced by Lamarck's famous *Philosophie zoologique* (1809). However, a manuscript by Lamarck survives with the title: "*Biologie/ou/Considérations sur la nature, les facultés, les développemens et l'origine des Corps vivans*"⁵. This is

thought to be a draft of a work on written between 1800 and 1801. In fact the word *biologie* was in use in 1800, before Treviranus and Lamarck. Karl Friedrich Burdach (1776–1847), anatomist and physiologist in Dorpat (today Tartu) and Königsberg, coined the word in his treatise: *Propädeutik zum Studium der gesammten Heilkunst. Ein Leitfaden akademischer Vorlesungen*, Leipzig 1800⁶.

But the very first time the word *biologie* appears in printed form is in the year 1797. In the *Vorrede* to his book: *Grundzüge von der Lehre von der Lebenskraft*, (Braunschweig 1797) Theodor Georg August Roose (1771–1803) referred to *biologie*⁷. Roose used the word only in the *Vorrede* and it does not reappear in the main text. This first edition of Roose's book was subject to a critical review, anonymously published in 1801 in *Neue allgemeine deutsche Bibliothek*, and the reviewer stated that Roose wanted to call "biology" what otherwise was named "physiology". So apart from Roose himself we have, in 1801, shortly after Burdach and slightly before Treviranus and Lamarck, another printed version of the word *biologie*.

Who was Th. G. A. Roose? Unfortunately, biographical material about Roose is scarce, but we know that he was a professor of anatomy in his native town of Braunschweig, and that he was awarded the title of *Hofrat* after not accepting a call to the university of Kiel in 1802. Apart from his own medical writings he was occupied with translating works by others such as those of Anton Scarpa from Pavia, Italy. He died, comparatively young, of a nervous fever.

GERHARD H. MULLER

Lehrstuhl für Biogeographie,
Universität des Saarlandes,
Saarbrücken, FRG

1. Magner, L.N. *A History of the Life Sciences* (New York, 1979).
2. Hoppe, B. *Colloque International "Lamarck"*, 199–273 (1971).
3. Morgue, R. *J. comp. Neurol.* 56, 503–510 (1932).
4. Stalleu, F.A. *Taxon* 20, 397–442 (1971).
5. Klein, M. *Archs Anat. histol. embryol.* 37, 105–114 (1954).
6. Schmid, G. *Nova Acta Leopoldina N.F.* 2, 597–620 (1935).
7. Dittrich, M. *Comm. Hist. Artis Med.* 83–74, 73–85 (1974).

Course geography

SIR — May I correct a geographical error in your article on Floyd Bloom's move from the Salk Institute to Scripps Clinic (*Nature* 3 March, p.3). As I remember the California coast at this point, Scripps overlooks the tenth fairway of Torrey Pines south course, and Salk runs alongside the 13th. Although my sense of direction (like that of my golf shots) is not outstanding, I figure this means that Bloom will in fact be moving northwards (away from Mexico), and not southwards as stated.

P.H. WILLIAMS

University of Leicester,
Leicester, UK

Cuba's great leap

Cuba may be moving into the big league of scientific nations at just the right time. Bob Ubell reports on a recent visit.

CUBA's leap into twentieth century science and engineering in only one generation has hurtled the island in most fields far beyond its Latin American neighbours. And in some applied areas — for example, sugarcane by-product research — Cuba has jumped to world leadership. Donald K. Luke, editor of the respected journal, *Sugar y Azúcar*, says, "The Cubans have the motivation and are now leading the pack."

Cubans admit that not all the paths they have taken from underdevelopment have led to success. But they point with pride to those roads which took them to high ground — including health care, cattle breeding and, significantly for their economy, sugarcane research. Their new push into more sophisticated terrain — computers, interferon, tissue culture and biotechnology — is not aimed at achieving "world class" results, as Eric Holtzman of Columbia University notes. Rather, "the primary value of basic research in Cuba is not that they are going to make major breakthroughs. It gives them a chance to familiarize themselves with the world literature and then to apply their knowledge to concrete problems".

Science in Cuba is not yet 25 years old. Modern research began only after the guerrilla army marched into Havana in 1959. Before the "triumph of the revolution", those seeking advanced degrees in science and engineering had to leave the country to study in the United States or in Europe. There were no graduate science programmes. Others could earn medical degrees at Havana University, founded in 1728. It is one of the oldest institutions of higher learning in the Americas.

The new government, committed to rapid development, started almost immediately to build science and technology with few human and material resources and with even less experience. "Cuba's level of scientific knowledge and experience has jumped immensely in the last ten years", Holtzman continues. "They surprise you all the time with how much they know. Their steady progress is amazing." Holtzman, a cell biologist, has collaborated with Cuban scientists since 1971 and reported on their work in 1979¹.

Cubans recognized early that science cannot be created without a critical mass of educated people. Soon after the revolution, in 1961, the government launched its acclaimed "literacy campaign", sending thousands of schoolchildren into the countryside to teach people how to read and write. Cuba now claims literacy rates far surpassing most of Latin America and the rest of the under-developed world.

By 1962 the government had reorganized higher education and its Academy of

Sciences. Until the mid-1970s, Cuba pushed thousands of young people through the universities and medical schools. "We were looking for large numbers", confessed one university official. "We created our first wave of professionals without taking into account their scientific level."

"In the early days, we really had no professors", admitted Ernesto Mario Bravo, a biochemist at the University of Havana who, together with other trained scientists, created a "cadre" of student teachers. "At night, after their day-time studies, we would train several dozen of the most brilliant third-year students how to teach second-year students. Without this system, it would have been impossible." Now, says Bravo proudly, members of these cadres "are directors, heads of laboratories".

Fidel Castro's slogan hangs prominently in many laboratories: "The future of our country must of necessity be the future of men of science". Castro's support for science and technology cannot be underestimated. He and his inner circle, mostly university trained, recognized the key role that advanced research was to play, even at first when most of the country's energies went into keeping body and soul together. Foreign Minister Carlos Rafael Rodríguez, one of the leading figures in the Communist Party and close to Castro, commented, "In the beginning of the revolution, science had inordinate support from Castro".

Scientific degrees, modelled on the Soviet system, were first awarded in Cuba in 1976, with the creation of the Ministry of Higher Education. By the end of 1980, the ministry reported that there were more than 200,000 students with a teaching population of just over 10,000 at 39 universities². Before the revolution, there were just three universities in the country. A new campus in Oriente province brings the present total to 40.

Today at the School of Biology at the University of Havana, nearly 2,500 students are offered classical biology programmes. About 100 come from abroad, largely from other Latin American countries, but also from Western and Eastern Europe. Tuition is free for Cuban and foreign students alike; some receive stipends. Of the 179 teachers at the school, 26 are professors.

The biology faculty operates the soon-to-be-completed national Botanical Garden, under construction sixteen years. A 600-hectare facility, near Lenin Park on the outskirts of Havana, the garden is, according to US botanist, "One of the most impressive I have seen. It rivals Harvard's." The school also administers the Centre for Marine Research on the coast

and runs Cuba's Anthropological Museum.

"We have had some success with research at the school", said Mario Oliva Suárez, vice-dean. Some novel experiments on cholesterol and studies on aquaculture, among others, appear to be going on at a moderately advanced level. Laboratories display a few modest pieces of equipment; exceptions are a new isotope facility with Japanese instrumentation and a student laboratory fitted out with monitors, recent gifts from Hungary.

According to Suárez, applied research represents about 40–45 per cent of the school's activities. Research on so-called "fundamentally oriented" problems accounts for some 50 per cent of their work; 5 to 10 per cent is devoted to basic research. In Cuban terms, applied problems are those near to commercial or pilot-scale production. Fundamentally oriented work covers research on those areas for which a solution does not appear immediate.

Students and faculty working on problems requiring high-powered instruments often go outside the university to affiliated centres elsewhere, notably Cuba's principal biomedical facility, the National Centre for Scientific Research (CENIC). CENIC and other laboratories run by the Ministry of Higher Education appear to act as links between the universities and industry. Apart from their mission as centres of excellence, modelled largely on high quality Western European and US-style laboratories, they also offer graduate students advanced training and provide industry with expert advice. The high-prestige centres capture Cuba's star scientists doing the most sophisticated work, leaving the universities with less elevated research tasks. Apparently aware that investigations in the schools need strengthening, the University of Havana's biology faculty, for example, plans to increase the number of full-time research positions from the present twenty.

"Basic research for Cuba is a seeding ground to enable their scientists to talk to others abroad", comments Columbia's Holtzman. "They are not aiming at high-prestige, low-productivity work as is often done elsewhere." Holtzman, who has worked with a CENIC group investigating electrical properties of cell membranes during phagocytosis, praises their approach: "They've aimed very carefully, planned very carefully to develop a scientific enterprise which is not intended to ape world class research", but rather to focus on the country's immediate needs. Juan Kouri, CENIC's director, agrees. "Research at the centre is multidisciplinary, covering both basic and applied research. Its main objective is to solve Cuba's problems."

But not all Cuban scientists are pleased. Some argue that intensive, direct studies aimed at specific economic aims can backfire, closing-off serendipitous routes to the

fulfillment of the island's needs. A few campaign for more support for fundamental work.

Cubans acknowledge that investments in research are a strain on the country's limited resources. While material conditions in the country have vastly improved in the past few years, certain items — beef, pork, even shoes — are still rationed. Nevertheless, Foreign Minister Rafael said, "We have decided that no financial obstacle will interfere with our programme in science and technology. This means we must continue to postpone consumption."

A high-ranking official in the Ministry of Higher Education, however, contended that basic research is still a "luxury" the country can ill afford. Significantly, Cuba's current five-year objectives for science and technology place "the greatest emphasis on applied research and development projects — irrespective of guided and other basic research. . ."

Created in 1965 as a springboard to propel research into modern fields, CENIC is housed in futuristic structures built in 1967 with sweeping stone columns in the manicured suburbs of Havana. It runs four main branches — biomedicine, chemistry, bioengineering and electronics. Under the wing of the Ministry of Higher Education, with links to the universities, Cuba's top research centre started life with merely a dozen scientists, at first based in laboratories in the vacated homes of wealthy Cubans who fled the country in the wake of the revolution. Large numbers of trained professionals, including nearly half the nation's physicians, left in the late 1950s and early 1960s, further damaging Cuba's prospects for development.

"The first thing we had to do was to organize ourselves as scientists", Kouri said. "We had nothing here." Kouri indicated that, in addition to help from a handful of US scientists, among them Roy John of New York University's Medical Centre and Holtzman, they received assistance from others, particularly Soviet experts, but also Western and Eastern European researchers. Continuing to rely on advice from abroad, at CENIC and elsewhere, the Cubans show a remarkable knack for selecting the world's top authorities. They have a nose for the best.

"A new institution takes about ten years to develop, to train people, to get apparatus, and to have enough technicians with skills to repair and maintain equipment", Kouri remarked. Adds Foreign Minister Rafael, "Even at the first stage of the revolution we made investments without really understanding or having an appreciation of science and technology. And now we have CENIC, which may be a model for other investments in science."

CENIC has a staff of one thousand, 350 of whom are professionals. Approximately 100–150 students do advanced work there, returning to their home universities to earn

their degrees. Some 50–60 staff members from other institutions also participate in research. "We need more biologists, more chemists, more engineers", pleaded Kouri in near-perfect, American-accented English. He stressed that he was expressing his own views, not national policy. Kouri, educated in Cuba, also studied at CNRS in France. Many scientists at CENIC and at other Cuban laboratories have been trained abroad, mostly in France, Belgium, Canada, Scandinavia and other Western European countries. Of course, many have studied in Eastern Europe, largely in the Soviet Union. Most scientists speak respectable English. English is the second language after Spanish in Cuba's secondary schools. Russian has just been introduced at a fairly modest level.

Most recently, the centre has turned to sugarcane tissue culture and cell fusion, monoclonal antibody work on alpha foeto-protein and, among other things, is working with Roy John on electrical activity of the brain. Kouri reported that molecular and cell biologists at CENIC are now "reforming" to create a coordinated biotechnology group.

Cuba produces more than 100,000 tons of yeast per year, largely for animal feed, so scientists at CENIC are working on ways to improve and nutritionally enrich its essential amino acids. While they have not yet found a cloning vector for *Candida utilis*, Carlos Pascual, head of biochemistry, claims that his group has been successful in working out the biochemical pathway of methionine in *Candida*⁴. Pascual also claims that the group has successfully extracted lysine from yeast and that this method has already been introduced into Cuban industry. Lysine will be produced as dry cells for poultry feed.

Visiting CENIC and other recently opened centres, observers from abroad are struck by the lavish architecture — miles of polished local marble, landscaped grounds and, most surprising, unusually well-equipped laboratories. Expensive, up-to-the-minute instrumentation comes from the United Kingdom, France and Scandinavia, with Japanese apparatus

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

dominating. Surprisingly, little or none is from the Soviet Union. "You can see where all their foreign currency is going", lamented one disgruntled Canadian businessman hoping to sell Cuba some nonscientific equipment. CENIC, one of the most well-endowed centres, operates some quite high-powered tools — a microprobe scanning electron microscope, for instance — mostly from Japan.

At CENIC and elsewhere, confronted with a vast amount of foreign-made apparatus, scientists are forced to devise novel repair and maintenance procedures. Contracts with manufacturers abroad usually call for Cuban technicians to go to the supplier for extensive training, even as far away as Japan. Often, spare parts are purchased together with instruments. So CENIC's electronics department manufactures its own routine machines, including a handsomely packaged pH meter which they also make for sale to other laboratories. They also assemble an EEG monitor, two of which have already been sold to Mexico. In-house training of technical support personnel, without foreign aid, is said to be quite rigorous.

The US embargo has had a serious and damaging effect on the progress of Cuban science, requiring laboratories to hunt elsewhere for reagents and equipment that could easily and cheaply be shipped from the United States, just 90 miles away. Forced to purchase materials from places far away from Cuba, scientists experience long delays and unwelcome additional shipping costs. Most troubling is the lack of current literature. While some libraries appear fairly up to date, other shelves look bare.

Yet the embargo has presented Cuba with unexpected benefits. It would never have developed its own highly organized and apparently efficient technical services, had it not been compelled. Since restriction enzymes, for example, produced principally in the United States, are unavailable, Cubans purify them themselves and, at an animal health facility, they make their own immunoassay kits. They have also built their own tissue-culture and virus collections. While much of this is tedious and un-



National Centre for Scientific Research
productive, Cubans acknowledge that some good experience is gained along the way and, most welcome, they earn a degree of independence.

Recognizing that poor access to the literature presents a serious roadblock, they claim to be taking extraordinary measures. By 1985-86, a new Central Library for Science and Technology will open with a stock of 7,000 periodicals and 10,000 books, selected from the world literature to match Cuba's research interests. The library will offer photocopying and other services aimed at circulating materials throughout Cuba. "Unfortunately", commented Humberto Arango, director of Cuba's Institute of Information Sciences, "we are not producers of scientific literature, we are consumers." A satellite link with the Soviet information retrieval system is soon to be hooked-up.

Cuba's science policy makers applaud CENIC's achievements; they have wisely used the centre as a base from which to launch other such groups. Rather than appealing to other institutions at, notably, the Cuban Academy of Sciences, planners have turned to productive and innovative CENIC scientists to spin off new laboratories elsewhere.

One of the most recent is the new Centre for Biological Research, also in Havana's suburbs, not far from CENIC. Much smaller, with a staff of 50, the unit was opened in January 1982. Mission-oriented, much like fledgling biotechnology firms in capitalist countries, it produces human leukocyte alpha interferon, using methods devised by Kari Cantell of the Central Public Health Laboratory in Helsinki. Cantell assisted in organizing research at the Cuban centre. In association with hospitals and medical teams, the group runs clinical trials. According to research vice-director Luis Herrera, one study in which alpha interferon was administered to patients suffering from dengue fever during a 1981 epidemic "stopped the complications" associated with the disease, especially in children, when the antiviral agent was used in the first days of onset⁵.

Herrera, in his early thirties, represents

the best of Cuba's young scientists. He was awarded Cuba's highest honour, Hero of Labor, for his achievements. He and others like him, teenagers when the new government took power, now occupy high-level positions. What's more, Herrera, who is black, is not unique; other blacks hold key posts.

Herrera indicated that the centre's interferon studies are designed to accomplish two principal tasks — to help eradicate or control tropical microbial and viral disease and as a model for genetic engineering. Recombinant DNA techniques for producing interferon are still at an early stage at the centre. Herrera and others believe that their move into bioengineering is important, suggesting that such studies will help close the gap between Cuba and more technologically advanced nations.

An autonomous laboratory, the centre falls completely outside the university system, the academy and other research umbrellas. It was created by an entirely new Cuban institution, the "Biological Front", a high-level policy-making body, formed two years ago. Designed to speed up focused research, the front offers new laboratories direct support from the government, short-circuiting bureaucratic restrictions. Doubtless, bureaucracy has held up progress in Cuba, as it does everywhere, particularly in Socialist states. To Cuba's credit, it has devised inventive techniques — the Biological Front being the latest one — to cut through red tape.

The other major installation recently spun off from CENIC is a sprawling complex in the countryside, less than an hour's drive from Havana. The National Centre for Animal and Plant Health (Centro Nacional de Sanidad Agropecuaria) is set in the midst of vast farms where test crops are under cultivation. Laboratories and animal research hospitals work on three main objectives — to attack diseases of cattle, pigs and poultry, to breed animals and to study diseases of Cuba's primary crops, sugarcane, tobacco and citrus fruits. CENIC was responsible for these lines before the new agricultural facility was created.

The animal and plant centre, and CENIC before it, were in part responsible for the development of Cuba's new cattle industry which at first was principally designed for dairy and is now turning to intensive beef production as well. Elisa Aznar Garcia reported that before 1959, cattle were raised in a few isolated regions, notably in Camagüey. "But there was no cattle industry as such." The early work created a high milk-producing tropical cow, bred from Canadian Holstein and Indian Zebu. This F1 variety, designed in collaboration with the Agricultural Ministry's Institute of Genetics, is now being replaced by a new Cuban-bred generation, known as Mambi. This stock produced "White Udder," a cow famous in Cuba for its astonishing milk production. Today, all milk and beef consumed in Cuba comes

from Cuban animals. But beef is still rationed.

In 1971, after some half-dozen years of pig breeding, just as the infant industry was about to take root, an African swine influenza epidemic ravaged the animals. The entire population had to be sacrificed, setting Cuba back five to seven years. Again, just as the next generation of Canadian pigs was ready for market, in 1981 another outbreak occurred. But because of experience gained with the first wave, the second outbreak was contained. But all the animals in the province of Havana had to be slaughtered; others were spared. Scientists at the centre failed to establish a link between the two outbreaks, nor could they find a "natural" cause for either occurrence. Swine flu is not native to Cuba. The government claimed that the introduction of the disease into Cuba appeared suspiciously like sabotage.

The centre's main line of work today is the development of vaccines against tropical animal diseases. A vaccine against cow diarrhoea has recently been created and is now in commercial production. Vaccines for other diseases are under study, including those against *Pasteurella*, influenza and *Salmonella*, in addition to one for *Escherichia coli* in swine.

Oddly, Cuba's top research groups are not under academy administration. They are run, instead, by the Ministry of Higher Education or autonomously under the new Biological Front. Cubans, unlike their Soviet friends, have chosen to put their research eggs into different baskets. In the Soviet Union, conservative academy institutes dominate research. But even in the Soviet Union, strict reliance on the academy is no longer the rule and newer, more innovative laboratories are emerging outside. When asked why a foreign science delegation was not shown even a single academy institute, an academy official smiled, "we have shown you our best."

Undoubtedly, Cuba's experience in quickly establishing a science enterprise from scratch has had much to do with this trend. After training skilled cadres in the universities, it must have seemed right to link the new research centres to the schools. Soviet academy institutes have long been criticized for their isolation from students. The many scientists who grew up in cadres and who were sent off to Western Europe for advanced study, must have been influenced by more liberal models.

It is also important that the island lies in the cultural shadow of the United States. In the streets of Havana, jeans, jogging shoes and tee shirts are ubiquitous. The Cubans are no less fashion conscious in their laboratories. Castro is said to have urged Cubans to study US textbooks, insisting that they are the best. While it may be surprising for visitors to find such evidence of US penetration in a socialist country, Cubans do not appear troubled. Rather, they welcome contact, seeking unofficial exchanges with US scientists despite the



The new Centre for Biological Research, Havana

official embargo.

Against this pattern, it is puzzling to discover that the highest level of Cuban science policy is in the hands of the academy. But when one finds that its current president is the man responsible for setting-up CENIC, things fall into place. Wilfredo Torres, a haematologist whose main research has been in sickle-cell anaemia, headed CENIC for twelve years from its inception. Rewarded for innovative administration, he now shepherds Cuba's entire science establishment. Torres not only runs the academy, but he is responsible for the direction and implementation of all research. His rank is equivalent to minister of state. Torres reports directly to José R. Fernández, Minister of Education and vice president of the Council of Ministers.

The Cuban academy dates back to 1861, but it was not until the 1960s that the new government reformed its structure. When Torres was appointed president in 1976, it was again reorganized. And as late as 1980, when all science-policy bodies were merged under its umbrella, it assumed its present form. Medicine, however, is run by a separate authority, the Ministry of Public Health (MinSaP). In 1962, the academy moved to its grand headquarters in the "Capitolio," a nineteenth century building modelled after the US Capitol in Washington, DC.

Of the 120 scientific, industrial and medical research institutes in the country, the academy operates 22. Cuban observers acknowledge that many of the academy's research facilities are largely inactive or working at a modest levels. They represent past legacies or the notion that each major discipline should have its own institute.

The academy employs approximately 1,000 research workers, or less than 10 per cent of the total number of scientists and engineers in the country. The academy also runs key national services, including the weather bureau, eight museums, three zoos and an aquarium, presumably accounting for a large share of its professional staff. Official figures show that in 1980, when the Cuban population stood at 9.9 million,

there were 11,400 researchers and 9,100 technicians. By 1985, Cuba expects to have more than 17,000 scientists and engineers and 12,500 technical personnel, out of an estimated 10.6 million people⁶.

The academy's principal function is as the country's highest-level science planning authority, especially since Torres took command. The Superior Scientific Council, composed of 77 distinguished elected scientists from the academy's institutes and others from the Ministry of Higher Education, industry and elsewhere, is one of two advisory bodies. A smaller, and very likely the more powerful advisory group, is the administrative council whose members are nominated by Torres and approved by Cuba's Council of Ministers. What's more, Torres relies on a team of close personal advisers — Eduardo Muzo, an MD, Gisela Alonzo, a chemist and plant physiologist, and José Carlos García, a veterinarian.

Cuba's research priorities are set by these groups, in collaboration with government and Communist Party officials who determine the main lines of work and allocate funds. Large investment decisions are made by the government's Central Board, the National Bank and the State Committee for Statistics.

Government officials outside the academy's policy-making bodies keep a close watch on research. Fidel Castro, like other chiefs of state, maintains a circle of personal advisers, including these who keep an eye on science. Castro's son, a nuclear physicist, is said to be one of them. Investigative commissions spawned by the National Assembly and the Communist Party also monitor research. The party's current five-year guidelines set out 17 priorities for science and technology, another eight goals for the "protection of the environment and natural resources" and hundreds of targets for industrial research and development.

One notable objective seeks to push forward the young computer industry now being developed by the autonomous National Institute for Automatic Systems and Computer Techniques. Headed by Daniel

Legrá, the institute is run inventively as a series of self-supporting "enterprises" selling services and equipment. Its mission is the total management of Cuba's computer network and the production, servicing and commercialization of computers for domestic and export sale.

As part of COMECON's computer system — in which each country is responsible for manufacturing a different piece of equipment for all members — Cuba produces a sophisticated alphanumeric, semi-graphic display unit. Designed especially for use with minicomputers, the units, assembled largely with Soviet components, went into production last year. In 1970, Cuba introduced its first mini. The Cubans are developing a "mega-mini" and now on the market is their third generation mini offering 64K bits with two bits per word. Foreign Minister Rafael remarked slyly, "Our friends in socialist countries counselled us not to invade this area. But we said we are tropical people and we are quick."

Obviously, Cuba is not going places without a little help from its friends. There is no doubt that the Soviet Union pours millions into the country. Russian tankers berthed in Havana harbour are the most visible evidence. But in 1980, Cuba claims to have invested close to £100 million of its own in research and development⁸, about one per cent of its national budget. This is a modest figure in comparison with most industrialized nations, but handsome for a developing country of Cuba's size. Little additional support comes from international agencies or Western European and Canadian research foundations. In 1980, only \$3.6 million came from outside sources, but grants from abroad are on the rise.

Cuba is not the Japan of the Caribbean. But in the 1980s the skill and organization required to move the island into high-technology does not take crippling investments of capital and labour. Bioengineering and digital computers do not come close to requiring what was needed in the nineteenth century to build Germany's chemical industry or install steel mills in the United States. Cuba possesses some key resources to enter the high-tech world — the wit to have trained a highly skilled population, the will to organize itself and the wisdom to choose the right targets. Fortuitously for Cuba, its entrance into advanced science and technology seems to have come at just the right moment. □

1. Holtzman, E. & Guttmacher, S. *Sciences* 24-27 (November 1979).
2. Leiner, M. J. *Reading*, 204 (December, 1981).
3. *Socioeconomic Guidelines for the 1981-85 Period*. 53 (Communist Party of Cuba, Havana, 1981).
4. Benítez, J. et al. *Folia Microbiol.* (in the press).
5. Ubell, R. *Bio/Technology* (in the press).
6. Fernández, D.O. & Sáenz, T.W. *El Estado Actual y Last Tendencias de la Política Científica y Tecnológica en la República de Cuba*, 39 (Academia de Ciencias de Cuba, Havana, 1981).
7. González, R.J. & González, G.M. *Electrónica y Proceso de Datos*, 1, (5), 82 (1982).
8. Fernández D.O. & Sáenz, T.W. *Electrónica y Proceso* 1, (5), 82 (1982).

No need for panic about AIDS

Acquired immune deficiency disease, now frequent among male homosexuals in the United States, is not this century's Black Death. The most urgent need is to understand what is going on.

THERE is now a serious danger that alarm about the disease physicians call acquired immune deficiency syndrome (unhelpfully, AIDS for short) will get out of hand. For the characteristics of this previously unrecognized and perhaps non-existent condition are so alarming that the temptation to portray it as a disease invited by a decadent civilization — a kind of latter-day version of the fate of Sodom and Gomorrah — is almost irresistible. Television companies are in danger of succumbing like flies to the temptation, for which reason it is important that the nature of the social problem that has arisen should be quantitatively as well as qualitatively understood.

Briefly, the condition is disturbing because it is either new or newly recognized, and because the mortality among those affected is high, probably more than 70 per cent. But the numbers of people among whom the disease has been recognized in the past few years is of the order of a thousand, most of whom are male homosexuals. Although successive waves of the mediaeval plague took several decades to work their way around the world, the communities in which infection became established would have been in no doubt (if they had known about the germ theory of disease) that some novel infectious agent had reached them.

The occurrence of acquired immune deficiency is, by comparison, covert. The first reports that something novel was afoot came in June 1981, and were based on the observation by the Center for Disease Control at Atlanta, Georgia, that five male homosexuals in Los Angeles had been diagnosed in the previous eighteen months as suffering from pneumonia caused by *Pneumocystis carinii*, a protozoan infection rarely seen by clinicians except in patients receiving immunosuppressive therapy. With the benefit of hindsight, the Atlanta centre was able to find further cases of this rarely seen pneumonia and also a geographically related cluster of patients who suffered or had died from Kaposi's sarcoma, a rare skin malignancy recognized at least five years ago as an accompaniment of immunosuppressive treatment (see Gange, R.W. and Jones, E.W., *Clin. exp. Dermatol.* 3, 135; 1978). The hunt for acquired immune deficiency disease, and for an explanation of it, has been on since then.

At this still early stage, it is too soon to know the scale of the problem, or even how

long it has been with us. Even if the causative agent is some transmissible agent, the fact that those who contract the disease turn up with some well-known but unrelated infection is bound to complicate diagnosis, which implies that the incidence of the disease is probably even now underestimated. Equally, however, since physicians on the look-out for a newly reported condition are less likely than in previous years to overlook the significance of an unusual condition, there is no easy way of telling how many people acquired immune deficiency before Atlanta recognized the importance of the first five cases reported from Los Angeles.

Indeed, the experience of the past two years, since the first reports in June 1981, suggests that there are many more than 1,000 cases to be identified. Among those so far found to be affected, male homosexuals predominate but acquired immune deficiency has also been found among women whose sexual partners are male homosexuals. Moreover, there are at least two cases in which newborn children have acquired the condition, one apparently *in utero* and the other as a result of blood transfusion. It is now recognized that the disease is unusually uncommon among those taking drugs intravenously, among their sexual partners, among those (such as haemophiliacs) dependent on blood products and — most surprisingly among Haitian immigrants to the United States without any of the predisposing dependencies of those among the indigenous population who succumb. But, so far as is known, the results of a study of immune deficiency in Haiti have not yet been reported.

Even the notion that immune deficiency is acquired by means of a transmissible agent, a virus for example, is open to dispute, or at least unproven. The evidence is only circumstantial, while the recent report (Ewing, E.P. *et al.*, *New Eng. J. Med.*, 7 April) of the presence of structures called vesicular rosettes in electron micrographs of the cytoplasm of cells from affected patients but only rarely in that from the cells of normal people is inconclusive, for the rosettes resemble no known virus. On balance, however, the epidemiology of the disease, and the observation that those dependent on blood products are especially at risk, suggests that something carried in the blood of affected people must be the transmitting agent. The peculiar susceptibility of Haitian immigrants to the United States who are apparently neither homo-

sexuals nor drug takers suggests some genetic predisposition.

When everybody is talking about parts of the human genome that may be mobilized to form viruses, it is natural that people's imaginations should take in the notion that integrable viruses such as human T-cell leukaemia virus (see *Nature*, 14 April) may be involved. But the presence of antibodies against such a virus may simply be a consequence of the impaired cell-mediated immunity of those with acquired immune deficiency. As it happens, male homosexuals not yet stricken with acquired immune disease tend also to carry antibodies against other viruses, hepatitis-B virus, for example, but that, too, is an inconclusive observation. Even the link between cytomegalovirus and Kaposi's sarcoma is inconclusive for the same reasons.

What, in the circumstances, should be done? Finding the cause of the disease is the most urgent need but, fortunately, the techniques are now available. The hope of treating those with acquired immune deficiency must be a matter of chance — an interferon or some other lymphokine may work, but only by sheer luck. Even the observation that, in many patients with acquired immune deficiency, the ratio of "helper" to "suppressor" T-cells is decreased is of no help when so little is known of the functions of these different types of cells in real life, or of their genetic origins. But mercifully, the disease — whatever its causation — is neither especially infectious (sexual contact seems necessary) nor certain in its effects. And the incubation period of a virus must be measured in months, not days. So there is a chance that the disease, whatever it is, will be understood before it has become too widely established, as bubonic plague once was.

Meanwhile, and for strictly prophylactic purposes, male homosexuals should be persuaded to change their ways. By all accounts, fears of acquired immune deficiency have already had a profound effect in cities where homosexuality is frequent. The pathetic promiscuity of male homosexuals is the most obvious threat to public health, but is probably no more serious now than it was before homosexuality ceased to be illegal. The use of substances such as amyl nitrite to assist the mechanics of male homosexuality, but which appear by accident to be immunosuppressants as well, should be discouraged. The fear of AIDS will no doubt help.

Tumour promoters

Protein kinase, phospholipid and control of growth

from I. Bernard Weinstein

THE mechanism of action of tumour promoters has until recently been largely mysterious. Unlike the initiating carcinogens, they cannot on their own induce cancer: but they can produce tumours when administered together with an ordinarily sub-carcinogenic dose of an initiating carcinogen. By contrast with the initiating carcinogens, which act directly on cellular DNA, many tumour promoters act primarily on receptors associated with cell membranes; but it has been unclear how. Recent studies suggesting that, in common with other growth-controlling agents, they may act through a specific protein kinase have therefore aroused considerable interest.

High-affinity saturable receptors for 12-*O*-tetradecanoyl phorbol 13-acetate (TPA) and structurally related plant diterpenes, as well as the recently discovered indole alkaloid and polyacetate classes of tumour promoter¹⁻⁴, have been identified in species ranging from nematodes to humans and in tissues as diverse as the skin, lymphocytes and brain.

It is clear that several of the structural and functional changes in cell membranes induced by the phorbol ester tumour promoters are mediated directly at the level of the cell membrane, since they occur within seconds or minutes of exposure and are not blocked by inhibitors of RNA and protein synthesis^{1,2}. Presumably, the subsequent pleiotropic effects on growth and differentiation, some of which do require macromolecular synthesis, represent secondary effects triggered by the initial binding of the tumour promoters to the membrane-associated receptors^{1,2}.

Results reported by Castagna *et al.*⁵, and later from other laboratories⁶⁻⁸, now suggest that the phorbol ester tumour promoters act by binding to and activating a specific calcium- and lipid-dependent protein kinase, protein kinase C.

It is widely recognized that the post-translational phosphorylation of serine and threonine residues in specific cellular proteins has an important role in mediating the action of various hormones and growth factors. At least four classes of such protein kinase (all of which require Ca²⁺) are known: cyclic AMP-dependent, cyclic GMP-dependent, calmodulin-dependent and phospholipid-dependent (protein kinase C)⁹. Certain oncogenes code for protein kinases that phosphorylate tyrosine residues; and tyrosine kinase activity is associated with the receptors for epidermal growth factor, insulin and platelet-derived growth factor¹⁰.

The properties of protein kinase C are

particularly intriguing, since this enzyme is ubiquitous in eukaryotes, it is strictly dependent on phospholipid (typically phosphatidylserine) and Ca²⁺, and its activity is markedly enhanced by unsaturated diacylglycerol^{5,9,11}. The latter findings suggest that the activity of this enzyme *in vivo* might be modulated by alterations in membrane structure and phospholipid metabolism, and thus it could play a key part in transmitting signals across membranes.

Castagna *et al.*⁵ have found that low concentrations (about 10⁻⁸ M) of TPA can substitute for diacylglycerol in the activation *in vitro* of rat brain protein kinase C, apparently by enhancing the interaction of the enzyme with Ca²⁺ and phospholipid. Phorbol compounds that lack tumour-promoting activity are inactive and TPA does not activate the enzyme in the absence of phospholipid. Evidence has been obtained that TPA also enhances the activity of protein kinase C in intact platelets⁵.

Niedel *et al.*⁶ have confirmed the finding that TPA activates rat brain protein kinase C activity *in vitro*. Furthermore, using divalent ion chelation they solubilized a protein from rat brain membranes which co-purified with protein kinase C, and bound labelled phorbol ester, but only in the presence of added phospholipid. Other investigators^{8,12} have detected a soluble protein in various murine tissues that binds tumour-promoting phorbol esters with high affinity and requires Ca²⁺ and phospholipid for maximum activity. The level of this protein in several murine tissues roughly parallels that of protein kinase C activity⁸. Curiously, the levels of phorbol ester-binding activity and protein kinase C activity are highest in brain.

It would appear, therefore, that protein kinase C may itself be a phorbol ester receptor, or at least a component of a receptor complex. Since, however, in the above studies neither the receptor nor the protein kinase was purified to homogeneity, further studies are required to establish that they are identical molecules.

TPA synergizes with, but does not substitute for, phospholipid in enhancing the activity of protein kinase C, and there is evidence that phorbol esters can bind with high affinity to synthetic lipid membranes and alter their physical properties, even in the absence of protein^{13,14}. Despite the fact that their chemical structures are quite different, the tumour-promoting phorbol esters, the indole alkaloid teleocidin and the polyacetate compound aplysiatoxin have somewhat similar hydrophilic and

hydrophobic domains^{1,4,15}. Because of their amphipathic character they may, therefore, bind to specific domains in phospholipid matrices.

The resultant changes in lipid structure could, in the presence of Ca²⁺, enhance the binding of phospholipid to the protein kinase C apoprotein. This could explain why, when intact cells are treated with TPA, soluble protein kinase C molecules seem to be translocated from the cytosol to membranes⁷. The hydrophilic domains of these tumour promoters could also interact specifically with the protein kinase C apoprotein to enhance the formation of a quaternary complex between phospholipid, the protein kinase, the tumour promoter and Ca²⁺.

Exposure of cells to phorbol esters, teleocidin or aplysiatoxin induces increased turnover of membrane phospholipids, presumably because of the activation of phospholipases including phospholipase C^{1,4,15,16}. This could generate diacylglycerol, itself an activator of protein kinase C^{5,11}, thus further augmenting protein kinase C activity. In this sense the long-sought putative endogenous analogue of these tumour promoters¹ might be a diacylglycerol. Perhaps the well known effects of dietary lipid on carcinogenesis¹ also act by altering membrane properties that influence the activities of protein kinases. It will be of interest to determine whether tumour promoters might also act by modulating the activities of other enzyme systems that are lipid-dependent.

The hypothesis that certain tumour promoters act via lipid-dependent protein kinases provides a satisfying unity to current research on growth factors, chemical carcinogenesis and tumour virology. In all cases, changes in protein phosphorylation may play a key part, although the precise kinases and target proteins that become phosphorylated may differ considerably. □

I. Bernard Weinstein is at the Division of Environmental Sciences and Institute of Cancer Research, Columbia University, New York, New York 10032.

- Weinstein, I.B. *J. supramolec. Struct. cell. Biochem.* **17**, 99 (1981).
- Hecker, E. *et al.* (eds) *Cocarcinogenesis and Biological Effects of Tumor Promoters* (Raven, New York, 1982).
- Driedger, P.E. & Blumberg, P.M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 567 (1980).
- Fujiki, H. *et al.* *Gann* **73**, 497 (1982).
- Castagna, M. *et al.* *J. biol. Chem.* **257**, 7847 (1982).
- Niedel, J.E., Kuhn, L.J. & Vandenbark, G.R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 36 (1983).
- Kraft, A.S. & Anderson, W.B. *Nature* **301**, 621 (1983).
- Ashendel, C.L., Staller, J.M. & Boutwell, R.K. *Biochem. biophys. Res. Commun.* **111**, 340 (1983).
- Flockhart, D. & Corbin, J. *Crit. Rev. Biochem.* **13**, 33 (1982).
- Kolata, G. *Science* **219**, 377 (1983).
- Kuo, J.F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 7039 (1980).
- Leach, K.L., James, M.L. & Blumberg, P.M. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Deleers, M. & Malaisse, W.Y. *Cancer Lett.* **17**, 135 (1982).
- Tran, P.L. *et al.* *Biochim. biophys. Acta* **727**, 31 (1983).
- Horowitz, A.D. *et al.* *Cancer Res.* (in the press).
- Mufson, R.A., Okin, E. & Weinstein, I.B. *Carcinogenesis* **2**, 1095 (1981).

Gene regulation

Prokaryotic mechanisms in eukaryotes?

from Charles Yanofsky

In view of the fundamental structural differences between prokaryotes and eukaryotes one might expect their metabolic processes to be regulated in quite different ways. So it has come as a distinct surprise to find a region of the yeast genome¹ with features hauntingly reminiscent of a mechanism of gene regulation hitherto considered exclusive to prokaryotes — that of control by the attenuation of transcription.

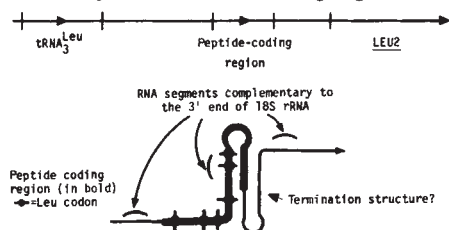
Attenuation of transcription is one of the ways in which the biosynthesis of amino acids can be regulated by the end-product of the synthetic pathway — the amino acids themselves. One example is to be found in the enzymes catalysing leucine synthesis in *Escherichia coli*. The genes encoding these enzymes are co-transcribed with an upstream sequence encoding a leader peptide rich in leucine residues². The control of leucine synthesis depends on alternative secondary structures adopted by the nascent transcript of the upstream leader sequence. When leucine is plentiful, the ribosomes read through the sequence encoding the leader peptide and the resultant transcript then adopts a secondary structure that terminates transcription upstream of the genes encoding the enzymes. In conditions of leucine shortage, however, ribosomes will stall at the leucine codons of the leader peptide, causing the stabilization of a secondary structure that prevents the termination of transcription and allows the downstream genes to be read³.

It now looks as though a similar mechanism may conceivably operate in the control of leucine synthesis in yeast (see the figure). Upstream of the yeast gene for the third enzyme of the leucine biosynthetic pathway (*LEU2*) there is a sequence potentially encoding a 23-residue peptide, six residues of which would be leucine. Moreover, sequences preceding *LEU2* could specify alternative base-paired secondary structures in the corresponding transcript — if indeed the region is transcribed. Particularly striking is a run of six Ts in the region immediately following one of these potential secondary structures — a feature typical of prokaryotic transcription termination sites.

The location of the potential secondary structures in the upstream message is such that if a ribosome stalled over a Leu codon in the peptide-coding region it could prevent formation of the secondary structure that resembles the prokaryotic transcription termination signal, exactly as is believed to occur in regulation by attenuation in prokaryotes³. The difficulty, of course, is that ribosome stalling could not occur in

the eukaryote nucleus because translation takes place only in the cytoplasm; so translation could not regulate transcription in the way it does in prokaryotes.

The analogy is further complicated by the fact that the 3' end of 18S rRNA, a segment that is often complementary to the 5'-non-coding region of yeast messages and may be involved in translation initiation, is complementary to the segment of the peptide-coding region that can form one of the secondary structures. The authors point out that the 100 or so nucleotide distance between the latter hypothetical ribosome-binding site and the AUG start codon of the *LEU2* coding region is within the range of distances observed in comparable regions preceding several yeast genes. However, alternative potential ribosome-binding sites are located just before the coding regions for



Organization of the *LEU2* upstream region and possible structures in the presumed combined transcript.

the Leu-rich peptide and the *LEU2* polypeptide. Also provocative but of unknown relevance is the existence of a gene for the major isoaccepting leucine tRNA of yeast, tRNA^{Leu}, immediately upstream from the peptide-coding region. This gene also has the same orientation as *LEU2* but is separated from the peptide-coding region by a likely RNA polymerase III transcription terminal site.

It is premature to attempt to accommodate these structural features in a comprehensive model in the absence of information on transcription and translation start sites and transcript-processing sites, and without knowing whether regulation is transcriptional, translational, or both. Nonetheless, the striking organizational resemblance of the *LEU2* upstream region to the regulatory regions of prokaryotic biosynthetic genes and operons has enticed the authors to speculation.

LEU2 of yeast is known to be coordinately regulated with *LEU1* and expression of both genes is controlled by the concentration of leucine or leucine plus threonine. These genes encode the proteins catalysing the third and second steps in the leucine pathway respectively. Expression of these genes is also influenced (induced)

by α -isopropylmalate (α -IPM), the product of the enzyme catalysing the first step in the leucine pathway. To accommodate the structure of the *LEU2* region with observations on its regulation the authors suggest that both the Leu-rich peptide and a complex formed between α -IPM and an inducer protein (for which there is some experimental evidence) may play a part in the control of transcription. They speculate that the Leu-rich peptide might act as a feedback regulator of *LEU2* expression. This peptide — synthesized when there is adequate leucine — could enter the nucleus and stabilize an RNA secondary structure that leads to premature termination of transcription. Under leucine deficiency — when the Leu-rich peptide would not be synthesized — premature termination of transcription would not occur and transcription would continue into the *LEU2* structural gene. Thus only when the leucine pool is large would enough leader peptide be formed to promote transcription termination. The negative effect of the peptide could be antagonized by positive interaction of the α -IPM-inducer protein complex with the transcript.

The authors further point out that translational control might also regulate *LEU2* expression. They suggest that the homology of the 3' end of yeast 18S rRNA with the segment of the Leu-peptide-coding region that is part of one of the potential alternative secondary structures implicates ribosome-binding activity in the control of *LEU2* expression. A possibly analogous mechanism exists in prokaryotes, in translational regulation of the synthesis of erythromycin methylase: erythromycin causes ribosome stalling in a short peptide-coding segment preceding the methylase-coding region, and alters a transcript secondary structure thereby activating the ribosome-binding site used for methylase synthesis⁴.

Alternatively, the Leu-rich peptide as such may have no regulatory function and transcription of *LEU2* of yeast may be exclusively regulated by the α -IPM-inducer protein complex. This complex, formed only under leucine deficiency, may act positively to stimulate synthesis of a long transcript containing the coding regions for both the Leu-rich peptide and the *LEU2* polypeptide. Initiation of translation of the *LEU2* segment of the transcript then may proceed only when the ribosome translating the Leu-peptide-coding region stalls at a Leu codon in response to a deficiency of charged tRNA^{Leu}. How this presumed stalling, the alternative secondary structures in the transcript and the three potential ribosome recognition sites preceding *LEU2* participate in the regulation of *LEU2* translation initiation awaits

1. Andreis, A., Hsu, Y.-P., Kohlaw, G.B. & Schimmel, P. *Cell* 31, 319 (1982).
2. Gemmill, R.M., Wessler, S.R., Keller, E.B. & Calvo, J.M. *Proc. natn. Acad. Sci. U.S.A.* 76, 4941 (1971).
3. Kolter, R. & Yanofsky, C. *A. Rev. Genet.* 16, 113 (1982).
4. Horinouchi, S. & Weisblum, B. *Proc. natn. Acad. Sci. U.S.A.* 77, 7079 (1980).

more definitive information on translation initiation in yeast.

Further speculation must await additional key findings. However it is already clear that the DNA region upstream from

Charles Yanofsky is at the Department of Biological Sciences, Stanford University, Stanford, California 94305.

LEU2 has structural features resembling those of the regulatory regions of prokaryotic biosynthetic operons regulated by attenuation. Thus further studies with *LEU2* of yeast may reveal how a regulatory region evolved from a prokaryote has become modified to allow regulation appropriate to a eukaryote. □

Stratigraphy

Is the present long enough to measure the past?

from Peter M. Sadler

AS students of geological history, stratigraphers must deal routinely with time spans that are far longer than their experience. From evidence preserved in layered sedimentary rocks, radical long-term changes in the configuration of oceans and continents, ice-caps and mountains can be charted. But the present, in which we can observe such processes in action, is incomparably brief. Geological processes are typically erratic, and their short-term characteristics can be very poor guides to their long-term behaviour. Stratigraphy can therefore sometimes lead to results that seem counter-intuitive in the light of present-day observations of geological processes in action.

Consider the magnitude of the time-scale discrepancy between the past and present in stratigraphy. The cumulative record of Earth history, as contained in stratified sedimentary rocks, spans three billion years. The shorter local sequences of rocks that constitute that record commonly span millions and tens of millions of years. Furthermore, it is still rare for time correlations between sequences to be more precise than one million years. In the oldest parts of the record, precision is far worse; within the deposits of the last one million years it can be far better; and of course, the present and historic past are studied with much greater precision still.

What is the time span of the present? Few individuals have studied one deposition system for even a decade. Continuous monitoring of sedimentation is rare, and seldom maintained beyond a few hours. Indeed, actualism in geology — the study of the present as a means to understand the record of past processes — has been common for little more than a century, following the leads of Hutton, Lyell and Darwin.

Yet sedimentology and stratigraphy deal with quite incongruous time scales. The incompatibility comes sharply into focus when modern rates of sediment deposition are used to estimate the time span represented by a sequence of analogous rocks. When this time span is known from radiometric dates or fossils, however, it is almost always found that the modern analogue seriously underestimates it. Thus modern

sedimentation rates are typically much higher than the accumulation rates of ancient rock sequences. Indeed, accumulation rates are known to decrease systematically as a function of time span in stratigraphical sections. The explanation appears to be that the strata of ancient rock reflect phases of erosion as well as deposition, and modern sedimentation rates may best be used as a guide to the time represented by a given layer of preserved rock. Thus the time span of specific layers expressed as a fraction of the total time span of the section gives a measure of the completeness of the record.

To the few empirical stratigraphical studies of time scale relationships, John Tipper (*Nature* 302, 696; 1983) has just added a very attractive statistical analysis, which integrates short-term sedimentological parameters and long-term stratigraphical phenomena. The attractiveness of Tipper's analysis is that it builds on a simple base that is consistent with present experience. He proposes to specify a sedimentation system in terms of the frequency distribution of its short-term rates of deposition. Negative rates represent erosion; and the mean, or expected, rate will be positive for a depositional regime and negative for an erosional regime. Longer-term accumulation can be modelled very simply from the frequency distribution by a random walk procedure, but it was not Tipper's objective merely to generate synthetic stratigraphical sections. He explores the general properties of longer-term accumulations.

The first result is that the expected long-term deposition rate is the same as the mean short-term rate, and not a function of time span. The empirical stratigraphical results appear to conflict with this finding, but this is because the two concepts of mean rate are not strictly analogous. Stratigraphical sections are the result of net accumulation and so represent only positive deposition rates. It follows that stratigraphical information cannot properly describe sedimentation systems.

Tipper investigates the probability that a time interval will be represented by any sediment accumulation. This probability is shown to be independent of the time span

of the section, but to decrease significantly as the time interval becomes shorter. Thus, stratigraphical completeness should be expected to fall as greater precision is demanded.

Tipper's analysis also predicts the behaviour of stratigraphical completeness for a given level of precision. The level of precision is fixed by the time scale of the short-term sedimentation rates on which the model is based, and completeness is the fraction of the total time span of a section that is represented by the active deposition time of the preserved sediment increments. Thus defined, the expected value of completeness is shown to decrease systematically with increasing section duration.

The completeness of a stratigraphical section clearly limits the scope of interpretations that may justifiably be made from it, but is difficult to estimate because it is ultimately determined by a unique sedimentation regime. Can Tipper's analysis provide estimates of completeness from sedimentological information? Unfortunately, the time-scale factor indicates that it cannot, because if the short-term time units are short enough for modern observations, then the precision of the completeness estimate becomes far too fine for practical stratigraphy. Furthermore, the frequency distribution of sedimentation rates cannot be properly characterized unless the systems are observed for long enough to include rare events. This would mean integrating the evidence of stratigraphical sections, and Tipper has shown that the data are not compatible.

If we generalize the stratigraphical accumulation rates, however, the modern deposition rates can be made compatible by the simple expedient of ignoring negative values. Tipper's analysis now serves to indicate the properties of an acceptable estimator of completeness. The simple estimator of completeness, overall accumulation rate of a section divided by a shorter-term mean deposition rate for the same environment, meets the criterion that the expected value should fall when estimated at finer precision. It succeeds because the empirical accumulation rates vary inversely with time span. The time scale of the short-term rates fixes the precision.

Such estimates of completeness can lead to counter-intuitive results. For example, our short-term experience of the episodic nature of fluvial deposition and the more uniform conditions of the shallow sea floor suggests that the river environment will produce less complete sections. At time scales longer than 10,000 yr, however, the similarity of stratigraphical accumulation rates dictates that we should not expect any difference on the basis of environment. Presumably, subsidence is more significant to the long-term accumulation than the short-term sedimentological processes that we can see. □

Peter M. Sadler is in the Department of Earth Sciences, University of California, Riverside, California 92521.

Superluminal motions

Astronomers still puzzled

from Richard Porcas

TWELVE years ago the astronomical community was startled by the claim of US radioastronomers to have discovered 'faster-than-light' motions in quasars. Results of repeated observations using the then infant technique of very long-baseline interferometry (VLBI) had revealed an angular expansion rate in quasar 3C279 which, when converted to a velocity using its redshift and the Hubble law, turned out to be around ten times the speed of light, c . Recently a small group of astronomers assembled at Jodrell Bank* to discuss the current observational and theoretical status of our understanding of these motions.

There are now seven known radio sources that exhibit superluminal motions, ranging up to at least $10c$, in their compact cores — 3C120, BL Lac, 3C273, 3C279, 3C345, 3C179 and NRAO 140. Unwin (Caltech) and Walker (NRAO) reviewed the detailed observational properties of the effect. The last six years have brought a great increase in reliability of results, particularly because of the use of many-telescope VLBI arrays and the routine production of milliarc-second-scale maps. Some apparently common properties of the compact structures are the existence of a compact inverted-spectrum core component, usually variable, and one or more extended components, which separate from the core with superluminal velocities. The extended components 'age', becoming weaker with time, and more rapidly at high frequencies, thus producing a steepening spectrum. They move in a roughly collimated region referred to as a jet, which, after bending, extends to the arc second scale. 3C345 has two such components, possibly separating from the core at different velocities. For 3C120 there are even more components, with different velocities.

There is some uncertainty about the constancy of the motion. Indirect evidence suggests that the outer component of 3C345 has always moved along its present position angle (although possibly accelerating) and that subsequent components emerge in different directions; even more recent data, however, suggest a curved motion for a new third component. Nor is it clear whether the compact core is indeed stationary (as is widely assumed) or whether it, too, moves — at present most VLBI observations only measure the relative motion between the core and other components.

Theoreticians do not believe that real

faster-than-light motions exist (thus breaking Einstein's speed limit) and most explain the effect as an illusion. Early on it was realized that a radio source moving with a relativistic (but subluminal) velocity can appear to move faster than light if the angle of the motion with respect to the observer's line of sight is small (because the observer's view of the early part of the motion is delayed with respect to later parts, and hence the apparent duration of the motion is shortened). Marscher (Boston University) reviewed the current theoretical models, concentrating on the relativistic jet model — a relativistically moving stream of plasma emanating from the core. The compact core of the superluminal sources is identified with the base of the jet and the 'moving' components are shocks or plasmions moving down the jet, which points towards the observer in the case of superluminal sources.

Some independent evidence exists for such relativistic motion. Observation of the X-ray emission from some radio sources yields fluxes many orders of magnitude less than would be predicted from the synchrotron turnover frequency and measured angular size, if the emission region were in the observer's rest frame. Lorentzian factors of the same order as those required to explain the superluminal motions are also sufficient to resolve this difficulty. In addition, they can explain the rapid variation in flux density of some sources at low frequency, reported by Fanti (University of Bologna).

Nearly all the superluminal sources also have arc-second-scale radio jets, on one side of the nucleus only, and a much-discussed question was whether this emission also comes from relativistically moving plasma, and whether the one-sidedness is intrinsic or caused by 'Doppler favouritism' of the approaching jet (that is, the flux density enhancement which occurs when relativistically moving emitting material moves at a small angle towards the observer). The same effect has traditionally been invoked to explain the relatively high incidence of the superluminal effect among the strong compact radio sources. Moore (Jodrell Bank) argued from statistical considerations that if such flux enhancement of the compact cores does indeed occur, then bulk relativistic velocities must also exist in the arc second jets.

Some investigations have been directed towards finding independent evidence for the rather small angles to the line of sight which supposedly exist in the superluminal sources. Samples of normal 'classical double' radio quasars should have a random distribution of source orientations, and VLBI observations of the compact cores of

such sources were reported by Zensus (Max-Planck-Institut für Radioastronomie, Bonn) and Barthel (University of Leiden). Porcas (Max-Planck-Institut für Radioastronomie, Bonn) reported on the superluminal behaviour in one such core in 3C179; the implied number of cores exhibiting superluminal motion may be uncomfortably large. Schilizzi (Netherlands Foundation for Radio Astronomy, Dwingeloo) presented maps of arc-second-scale resolution from the Westerbork array which, using a new high-dynamic-range technique, have revealed faint outer double structure in all but one (3C273) of the known superluminal sources. An embarrassment is that the average projected size of the outer structure is no smaller than that of the normal radio-source population; thus the outer structure apparently is not at the small angle to the line of sight required for the inner jet, and one must postulate a misalignment between inner and outer source structure.

The 'unified scheme' of Browne (Jodrell Bank) interprets all radio quasars as members of a single class, the properties of individual quasars depending on the orientation to the observer. Thus, sources with enhanced cores and arc-second-scale jets are those closely aligned to the line of sight, whereas quasars with weak radio cores and outer double lobes have a large angle to the line of sight. Some properties of radio sources are indeed consistent with this scheme; sources with strong cores tend to have bent outer structures (the small angle to the line of sight would magnify any intrinsic small bend by projection), which have, on average, a smaller angular size (foreshortened by projection). Other properties are less readily explained; for example, the jet in double-lobed sources tends to point the lobe with the most compact or multiple hot spot (outer lobes, being unbeamed, should not betray whether they are on the near or far side of the nucleus).

Many of the problems in accepting the relativistic jet model arise from the small line-of-sight angle required. This problem can be eased if plasma emerges not in a single direction but in a 'spray' of directions which must later be recollimated. Scheuer (University of Cambridge) spoke of one version of this, the 'Blunderbuss', in which many ejecta are shot out in different directions, the observer preferentially seeing those making a favourable angle to his line of sight. This model, however, does not seem able to reproduce the narrow core-jet structure seen in the superluminal sources.

Blandford (Caltech), following the principle that there is 'no extragalactic smoke without fire', discussed current ideas about the central compact objects which produce the plasma beams. He introduced his 'grand unified scheme' in which the whole range of phenomena including the Galactic Centre, Seyfert galaxies and quasars are interpreted as reflecting just two intrinsic parameters — the mass of the central com-

*A 'Superluminal Workshop' was held at Jodrell Bank in February 1983.

compact object (black hole) and the accretion rate onto it.

In the final session, Scheuer led a discussion of some of the central questions. Is there bulk relativistic motion? If so, how much more complicated than the simple relativistic jet must it be to account for all the facts? What bends jets, are they intrinsically one-sided and how fast are the arc second jets? What is the relationship between VLBI jets and large-scale structures? What are the family relationships?

It seems that 12 years of intensive study of superluminal motion have raised more questions than they have answered. Perhaps the next 12 years, with the prospect of a dedicated VLBI array studying the structural changes in large numbers of objects, will bring answers to some of these questions. □

Richard Porcas is at the Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-5300, Bonn 1.

Schistosomiasis

Plants that kill snails and prospects for disease control

from G. Webbe and J.D.H. Lambert

MOLLUSCICIDES play an important part in the effective control of schistosomiasis. Partly as a consequence of the rising costs of synthetic molluscicides, there has been an increasing interest in plants showing molluscicidal properties. The plants *Phytolacca dodecandra*, *Ambrosia maritima* and *Anacardium occidentale* are among several which have shown promising molluscicidal activity. At present certain saponins, specifically those with an oleonic acid derivative, seem to offer the greatest potential as molluscicides.

A significant reduction of schistosomal disease may be possible with the use of the new efficient drugs. Although chemotherapy will undoubtedly play a vital part in reducing the severity of disease and contamination in the environment, as well as helping to control the transmission, the effect of population-based chemotherapy as the sole method of controlling transmission is unknown. It is however considered that, in many areas, a combination of snail control and chemotherapy will be the quickest and the most cost-effective way to reduce the prevalence and intensity of infection and to control transmission.

Even so, where schistosomiasis is not a clear priority, the high cost of drugs, molluscicides and other methods will clearly make it very difficult for poorer countries, even in major endemic areas, to establish control programmes within the constraints of small health budgets. Control programmes will not be seriously undertaken, nor will substantial progress be achieved in such areas, unless overall costs are generally less than \$1.0 per caput of the population protected. It is therefore important to maintain efforts to develop lower-cost technology including drugs, molluscicides and their application.

The impetus to develop plant molluscicides is largely due to a combination of the high cost of imported synthetic compounds and increasing concern about the possible development of resistance to them, and their toxicity to non-target organisms. Plant molluscicides will probably be most

useful in areas where transmission is predominantly focal, rather than widespread through large areas in which synthetic molluscicides will continue to be used. It is believed that plant molluscicides could play an important part in integrated programmes at the village level and in self-help schemes.

Since the 1930s, more than 1,000 plant species have been tested for molluscicidal activity, including nearly 600 in China, where a water extract of the oil of tea-cake (*Thea oleosa*) has been used as a molluscicide and plant fertilizer, and the powdered residue as a cercariacide in the footwear of farmers. The most thoroughly studied plant molluscicide is one derived from the berries of *P. dodecandra* (endod), an aqueous extract of which has been used for snail control in Ethiopia. In the Sudan, a water extract of the seed of *Croton macrostachys* has high molluscicidal activity, as has that of *Jatropha curcas*, while in Egypt trials are in progress using *A. maritima*. The crushed shells of the cashew nut (*A. occidentale*) are being used for snail control in Mozambique.

Ten other species of *Ambrosia*, which is essentially an American genus, contain ambrosia and/or damsin, the two sesquiterpene lactones isolated from *A. maritima*. Some 42 molluscicidal compounds have been isolated from plants, and in addition to saponins, thought until recently to be the major active molluscicidal ingredients, flavonols, alkaloids, terpenoids and other groups have produced encouraging results. The recent finding that bidesmotic saponins and some dammarane glycosides and alkaloid saponins are inactive may help to explain the variation in the results obtained by different investigators. Further studies of other structure-activity relationships of compounds may result in the identification of plants with more selective effects. Molluscicidal compounds, like other phytotoxins, may be present and synthesized more frequently in some plant groups than in others: thus, sesquiterpene lactones

are most common in the Compositae, although triterpenoid saponins are widespread in nature.

The relatively high proportion (58 per cent) of active species among the 491 species belonging to 105 families recently tested is apparently due to the selection of medicinal plants for primary screening by most investigators. The highest proportion of molluscicidal species is in the Families Phytolaccaceae followed by the Mimosaaceae, Polygonaceae, Verbenaceae and Caesalpiniaceae. A large proportion of Euphorbiaceae are also highly active.

There is little information on the distribution of the molluscicidal principles in different plant parts, which might have predictive value, owing to the failure of most investigators to study whole plants systematically. In the Phytolaccaceae species, the berries are most active, the leaves less so and the roots not at all, which agrees with the distribution of saponins and inactive alkaloids. In most species in the Family Euphorbiaceae, the highest activity is in the seeds, and in the Solanaceae, in the leaves and fruits. Thus while efforts should be made to test all parts of the plant for molluscicidal activity, clearly for sustained use in snail control programmes it will be best in the case of perennial plants to concentrate on regenerating leaves, fruits and flowers rather than on roots, stems and non-regenerating bark.

Little is known about snail metabolism or physiology which can be used to predict the structure of potential synthetic molluscicides, let alone substances of plant origin. The known active principles of plant molluscicides are water-soluble and act on the snail surface membrane, the dose-response of saponins being an all-or-none phenomenon over a very narrow range. In considering the biochemistry of recognized molluscicidal compounds of plant origin, saponins have been emphasized, but others should be considered including tannins, flavonoids, alkaloids and terpenoids.

The availability and accessibility of information on the geographical distribution and ecosystem of molluscicidal plants in America and Europe will be extremely valuable for future research. The use of toxic plants for agricultural as well as vector control purposes may be possible; and the frequent absence of aquatic snails in the presence of specific plants may suggest which plants may be suitable for biological control.

The cultivation of plants on a large scale exclusively for snail control has not yet been attempted. The agronomic success achieved in the 1970s with *P. dodecandra* would suggest that, at least in Ethiopia, cultivation at the village level is feasible. □

G. Webbe and J.D.H. Lambert are in the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases at the London School of Hygiene and Tropical Medicine, Winches Farm Field Station, St Albans, Herts.

Photoperiodism

Melatonin as a seasonal time-cue: a commercial story

from Gerald Lincoln

THE physiology of many of our domesticated animals varies with the seasons in the same way as that of their wild ancestors. Sheep, goats, horses, ferrets, mink and deer are seasonal in their reproduction, while cattle show a seasonal cycle in milk yield. In most animals, appetite increases in summer, fat reserves accumulate in autumn and a period of anorexia follows in the winter. Moulting also occurs seasonally, as does the growth of horns and the replacement of antlers. These seasonal traits all have obvious survival value in the wild, but in captivity they can be constraints on management; it is, for example, difficult to induce sheep or horses to conceive during their natural non-breeding season, or to make cattle and deer put on weight in the winter. One way to avoid this problem is to house the animals indoors and expose them to patterns of artificial light. A more convenient alternative, now being discussed, is to treat the animals with melatonin.

The new approach comes from the realization that melatonin, released from the pineal gland during the hours of darkness, relays the photoperiodic time-cues through which seasonal species respond to the annual cycle in daylength¹⁻⁵. The period for which melatonin is at a high level in the blood appears to convey 'night-length' and to cause specific changes in the hypothalamic control of reproduction, moulting, food intake and other seasonal characteristics. By giving animals melatonin it is possible to interfere with this subtle time-keeping mechanism. Increasing the length of time for which circulating melatonin is at a high level has been found to induce the short-daylength response normally seen in winter in all the mammalian species studied so far.

There can be important differences between species in the effect of the melatonin signal, however. Reproduction is variously affected: long exposure to melatonin stimulates gonadal activity in sheep, goats and deer, all of which breed when daylength is short; but inhibits gonadal activity in horses and many smaller species, including hamsters, ferrets, voles, weasels and hares, that breed when daylength is long. Effects on moulting and appetite are more uniform: melatonin causes growth of the winter coat and anorexia.

The addition of melatonin to the diet is now a practical method of treating farm animals. Dave Kennaway and Bob Seamark⁶, working in Adelaide, have shown that giving sheep 2 mg of melatonin absorbed onto a pelleted diet produces a physiological 'night-time' increase in the

blood levels of melatonin lasting up to 8 h, the length of the effect being a result of slow absorption from the gut. They subsequently used this technique to administer melatonin to anoestrus Merino cross-bred ewes housed in summer daylengths of 16 h light: 8 h darkness. Feed containing the melatonin was given daily 8 h before 'lights-off' on the ground that the increase in melatonin blood levels it produced would merge with the increase in endogenous melatonin secreted during the dark phase and expose the animal for long enough to induce a short-day response. The effect was as predicted: the high blood levels of prolactin, typical of long daylengths, decreased within 17 days, heralding an early resumption of the oestrous cycle. The effect has been confirmed by Jo Arendt⁷ in Gilford.

Oral administration of melatonin has also been found to be effective at inducing antler maturation and rutting behaviour in male white-tailed deer⁸. Melatonin can even be given in the drinking water to produce a short-day response⁹, or simply as a small subcutaneous implant which will release a constant level of melatonin over many weeks¹⁰⁻¹². One limitation is that prolonged treatment with melatonin leads to refractoriness¹³, as occurs during prolonged exposure to short days, and only by exposing animals to long days with no melatonin can responsiveness be restored.

Two models have been proposed to account for the actions of melatonin in photoperiodic time-measurement. In one, it is the total duration of melatonin exposure each day which dictates the response, while in the other it is the timing of a daily rhythm in sensitivity to melatonin relative to the period of melatonin secretion which dictates the response. In the first model, the duration of melatonin exposure is assessed in the brain by some form of hour-glass. In the second, the critical feature is the way the daily photoperiod entrains the phase of the rhythm in brain sensitivity and the rhythm in melatonin secretion.

Supporting the first hypothesis are the results of numerous experiments in which the duration of exposure to melatonin has been increased by the addition of exogenous melatonin to induce a short-day response. Elegant studies performed by Bruce Goldman^{14,15} working with Djungarian hamsters and Eric Bittman¹⁶ working with Suffolk sheep, in which programmed infusions of melatonin are given for a set duration each day, have provided compelling evidence for the importance of duration. If the blood levels of melatonin are continuously elevated for more than 12

h every day for several weeks, a short-day response is seen; but if they are elevated for only short periods, even if more than once each day, then a long-day response is seen.

These experiments do not, however, totally rule out the second 'time-of-day' hypothesis. Larry Tamark¹⁷ first showed in 1976 that injecting hamsters on long-daylight regimes with melatonin caused a short-day reproductive response if the injection was given in the subjective 'afternoon' but not if given in the 'morning'. However, the treated animals had functional pineal glands and when similar experiments were performed with pinealectomized hamsters no time effect was apparent^{18,19}.

The debate still continues, though, for in recent experiments, Marcia Watson-Whitmyre and Milton Stetson²⁰ show that two precisely timed injections of melatonin given to hamsters kept on different photoperiods produce different responses.

Finally, it should be noted that although the administration of melatonin is a convenient method of inducing winter changes during the summer, there is no simple way of achieving the reverse. To do this, we need a powerful antagonist to block the action of melatonin in the brain, or a compound which specifically inhibits the release of melatonin from the pineal gland. Armed with these anti-melatonin compounds it may be possible to induce an increase in appetite of farm animals in winter and to stimulate early breeding in long-day species. It might even be possible to induce race horses to foal close to 1 January: the aim of all horse breeders. □

Gerald Lincoln is at the MRC Reproductive Biology Unit, Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh EH3 9EW.

1. Wurtman, R.J., Axelrod, J. & Kelly, D.E. *The Pineal* (Academic, New York, 1968).
2. Wurtman, R.J. & Anton-Tay, F. *Recent Prog. Horm. Res.* **25**, 497 (1969).
3. Turek, F.W. & Campbell, C.S. *Biol. Reprod.* **20**, 32 (1979).
4. Reiter, R.J. *Endocr. Rev.* **1**, 109 (1980).
5. Herbert, J. in *Biological Clocks in Seasonal Reproductive Cycles* (eds Follett, B.K. & Wright D.E.) 261 (Wright, Bristol, 1981).
6. Kennaway, D.J., Gilmore, T.A. & Seamark, R.F. *Endocrinology* **110**, 1766 (1982).
7. Arendt, J., Symons, A.M., Laud, C.A. & Pryde, S. *Abstr. 58 Soc. Study of Fertility Conf.*, Nottingham (1982).
8. Bubenik, G. *J. exp. Zool.* **225**, 155 (1983).
9. Pevet, P. & Haldar-Misra, C. *Experientia* **38**, 1493 (1982).
10. Turek, F.W. *Proc. Soc. exp. Biol. Med.* **155**, 31 (1977).
11. Allain, D., Martinet, L. & Rougeot, J. in *Photoperiodism and Reproduction in Vertebrates* (eds Ortau, R., Pelleier, J. & Ravault, J.P.) 263, (INRA Service des Publications, France, 1981).
12. Lincoln, G.A. *Abstr. 57 Soc. Study of Fertility Conf.*, Nottingham: unpublished observations (1982).
13. Bittman, E.L. *Science* **202**, 648 (1978).
14. Carter, D.S. & Goldman, B.D. *Biol. Reprod.* **24**, 23A (1981).
15. Goldman, B.D. *Soc. Study of Reprod.* **26**, Abstr. 22, Suppl. 1 (1982).
16. Bittman, E.L., Karsch, F.J. & Dempsey, R.J. *64th a. Mtg Endocr. Soc.* (1982).
17. Tamarkin, L. *et al. Endocrinology* **99**, 1541 (1976).
18. Tamarkin, L. *et al. Science* **198**, 953 (1977).
19. Goldman, B. *et al. Endocrinology* **104**, 82 (1979).
20. Watson-Whitmyre, M. & Stetson, M.H. *Endocrinology* **112**, 763 (1983).

Astronomy

Galactic irregularities — nature or nurture

from Simon D.M. White

FOR many years after Hubble demonstrated the validity of Kant's view that galaxies are separate 'island universes', astronomers tended to treat galaxies as independent closed systems well removed from external influences. Their form was assumed to reflect the conditions in which they were born and the manner in which they had aged. Recent research, however, has placed ever greater emphasis on environmental effects and on interactions between galaxies and their neighbours. Such interactions seem likely to have left their imprint on many aspects of galaxy structure and offer a possible explanation for a number of irregularities often seen in the outer regions of spiral galaxies. Nevertheless, a recent paper by Krumm and Shane (*Astronomy and Astrophysics* **116**, 237; 1982) shows that the case for an external rather than an intrinsic source for these irregularities is still far from proven.

The authors made maps of the distribution of interstellar hydrogen gas in two of the most isolated and asocial spiral galaxies known; they found the galaxies to have marked asymmetries similar to those which in other systems have been ascribed to the influence of nearby companions. This evidence suggests that both 'nature' and 'nurture' need to be taken into account to explain the origin of the features.

The idea that environment could have a determining effect on galaxy structure stemmed from a number of observations. The global morphology of galaxies is clearly related to the proximity of neighbours. In regions of space where the density of galaxies is high, elliptical galaxies are common, spiral galaxies are rare and galaxies tend to have little interstellar gas. In lower-density regions, spirals form the dominant population and ellipticals are much rarer; most galaxies have a significant interstellar medium. This strong trend shows either that galaxies 'knew' what their current environment would be at the time they were formed, or that their current morphology is a result of environmental influences.

Such influences are directly observable in certain galaxy pairs that are just now in the process of colliding and that show strong distortions characteristic of tidal interaction. The time during which these distortions are readily apparent is quite short, so many apparently quiescent systems may have been subjected to violent distortions in the past. Permanent structural changes must have been induced by such distortions. The value of studying extremely isolated galaxies, such as those selected by Krumm and Shane, is that they provide a control group for which past

strong encounters seem very unlikely; there are just no candidate perturbers.

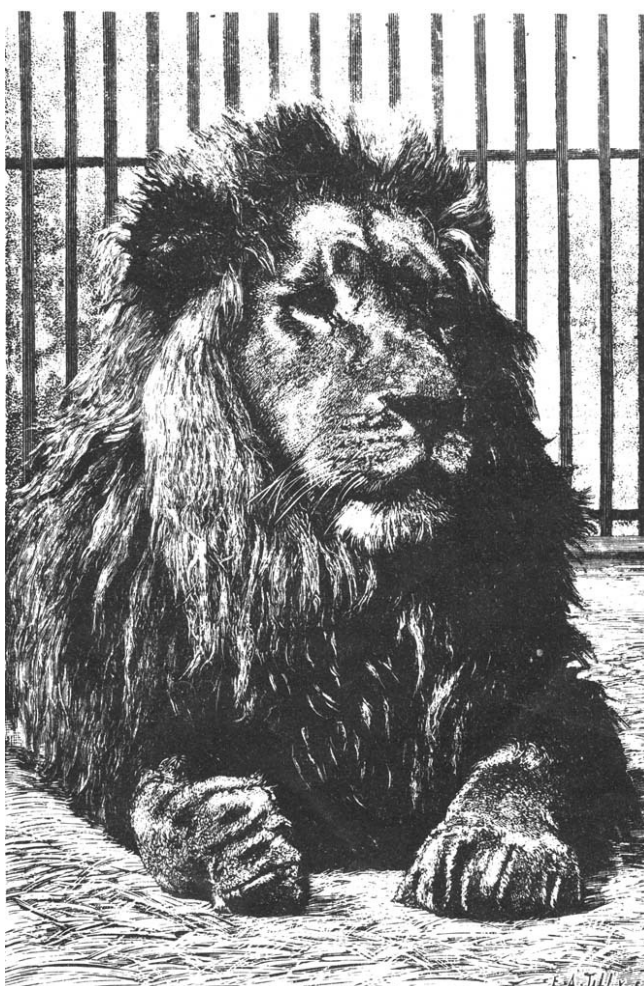
Several kinds of structural change may be induced in a spiral galaxy by encounters with other systems. The gravitational field of the intruder may rip off weakly bound material in the outer regions and thus tidally truncate the galaxy's mass distribution. At the same time the distortion of the remaining material may produce long-lived warps near the edge of the galaxy and may trigger the conversion of a significant amount of gas into stars. Recent studies have suggested that elliptical galaxies may indeed be tidally truncated in regions of high galaxy density, that warps and global spiral patterns in spiral galaxies can be related to the presence of companions and that galaxies in close pairs have higher rates of star formation than similar isolated systems.

These trends are at best only rather weak, however, and striking exceptions can be found. Despite their isolation the two galaxies of Krumm and Shane are no

more extended and contain little more gas than typical spirals. Moreover, one of them shows a pronounced warping of its hydrogen disc, and the hydrogen distribution of the other is both locally clumpy and globally asymmetric. In these two cases, at least the asymmetries must be intrinsic and must reflect either internal instabilities or a very slow rate of settling to a symmetric state.

Neither possibility fits well the standard models for the structure of galaxies, and at an IAU symposium in Besançon last summer A. Toomre concluded that there is at present no convincing theoretical explanation for the existence of warped discs in spirals which do not have companions. His tentative best bet was that the observed galaxy is surrounded by a massive but invisible halo which is flattened along an axis inclined to that of the central disc. In the outer regions, the disc then bends towards the principal plane of the halo. This model amounts to replacing the asymmetric gravitational field of an observed companion by the asymmetric field of an unseen halo when no such companion is in evidence. It accounts quite prettily for the observations, but there is so far no other evidence for such a misaligned dark halo. □

Simon D.M. White is in the Berkeley Astronomy Department, University of California, Berkeley, California 94720.



100 years ago

THE LION AT REST

THIS illustration from *La Nature* is after a photograph of one of the lions in the Zoological Gardens, London. This photograph may be regarded as one of the numerous triumphs of instantaneous photography, valuable both to art and science. The original was rephotographed in Paris directly on wood, by means of a special collodion, at present much used. This has assured a perfectly faithful reproduction of the original, exhibiting all the characteristic details of the lion at rest. From *Nature* **27**, 584, 19 April 1883.

The book review business

Most science journals include book review pages. Tim Lincoln discusses what there is in the book review for a book's publisher and author, the reviewer, and the journal and its readers.

REPORTS of the death of the book have been exaggerated, not least in science. Last year in Britain over 48,000 books were published. Of these 11,000 came into the category of science, technology and medicine (STM).

Some of the new books among the 11,000 were no doubt dross, some jewels. Those judgements, and others in between, will in due course be confirmed by low or high sales and perhaps the eventual compliment of a second edition. More immediately they are to be found in the review pages of periodicals, science journals among them. What is there in the book review business for the four principals concerned — the publisher; the author or editor of a book; the reviewer; and the journal and its readers?

Of the four, it is publishers who would seem to have the clearest reason for wanting to see several inches of (preferably prompt and laudatory) comment upon one of their books. Publishers, the cynic would say, are simply in the business of selling a product and that publishers seek reviews because reviews sell books. Like everyone else the cynic is guessing. No one knows whether good reviews sell books, whether bad ones stop them selling, or indeed whether reviews affect sales much at all. Trade publishing mythology speaks of "selling reviews" by the likes of Arnold Bennett and Bernard Levin, but it seems unlikely that a review of Volume 13 of *Advances in Toluene Research*, appearing a year after the book's date of publication, will bring a glint to librarians' eyes and their hands to the cash box. Besides, the biggest STM publishers have such a tight grasp of their mailing lists that the review as a vehicle for selling may be superfluous.

Sending out review copies is expensive. Given the ratio of review space to books published, many of them gather dust in editorial offices and may eventually be interred on the shelves of a second-hand bookshop. Reviews must surely help the sales of some books; and they at least advertise the subject areas in which a publisher is active. But from the publishers' viewpoint there are further reasons why the review ritual continues. First, perhaps, simply because it is a ritual. More positively, publishers genuinely and justly look for constructive criticism of their "product" from as many impartial sources as possible. (That is why it is such a dreadful sin for a reviewer to pronounce on the same book in different publications; review space is hard won.) For publishers, however, perhaps the most compelling

reason for seeking reviews is that an author without a review is an unhappy author.

Cynicism apart, publishers need to make a profit. They like also to publish respectable books that are well-received in public. Conversely academic authors also like to make money but their greater need may be to write good books. Perhaps a few colleagues will have commented upon some of the chapters at manuscript stage or may even have written some of them. Whatever the form of the book, however, or the motive behind writing or compiling it, it is understandable that any author or editor should wonder "How did I do?". The corollary is that the question should be echoed and answered in print in the form of a review. There are to be sure a few books which nobody seems to care about, the symposium proceedings' volume usually being among them. But more often an author has put not a little bit of his or her soul into the writing of a book. In this respect it is a solemn task that reviewers have pressed upon them, though one of which authors historically have taken a poor view.

Reviewers, according to Shelley, "with some rare exceptions are a most stupid and malignant race". Samuel Coleridge was more scornful: "Reviewers are usually people who would have been poets, historians, biographers. . . if they could". Scientists as book reviewers are usually spared such libels, the harsh comments applying not to all of the diverse forms of reviewing but to arts' criticism and most especially to professional critics. Snug in peer review, science and the STM book lack such critics. Even though the peer book review too is the norm — for the good reason that STM books are so technical in content and language — it is nonetheless here that there is a happy oddity in the practice of science.

Custom decrees that in book reviews it is usual for a reviewer to put his name publicly to an evaluation of the work of others. By contrast, a different tradition dictates that much of the rest of criticism in science should be anonymous, for example in the appraisal of grant applications and refereeing of papers. It is true that there are iconoclasts in the second respect, yet in being signed the book review is an anomaly. It is at the same time a respected anomaly, so that when a journal invites someone to pontificate in print on a book, that person will often agree. But why?

The most conspicuous motives for taking on the chore of a book review — the mercenary and the self-serving — are also

the most superficial. There is the lure of possessing the review copy, the standard incentive and often quite a prize when the £50 academic book is not uncommon. There is also the opportunity of self-publicity, or of riding a hobby-horse. And there may be a fee, though reviewing academic books even less than writing them is financially unrewarding. All of these may weigh in the decision. Yet there is a more general and powerful reason. Most often reviewers commit themselves to read a book, and write about it, as an extension of their academic duties and as an exercise in self-education. The imposed discipline of copy deadline and word limit may or may not be a necessary goad. The result is nevertheless that some-of-those-books-which-ought-to-be-read are read, and *read* not merely their pages turned. The beneficiary of such labour is not only the reviewer; whether in exposing delusion or unearthing treasure, the journal and its readers are also served.

All of this is the earnest stuff of scholarship, but at its best a review earns space in the books' pages not merely by cataloguing the chapters in the volume and critically evaluating them, but through the extra quality of general interest and the intangible merit of "readability". Quite obviously a book such as Lewis Thomas's *The Youngest Science* (reviewed on p. 778), about which there is a story to be told, allows that third function to be taken further than the apocryphal Volume 13 of *Advances in Toluene Research*. Equally, greater discursiveness is made possible by a limit of 2,000 rather than 200 words. Nonetheless that a brief review of a scholarly book can be illuminated by wit, learning and pungent asides is shown in the review pages of a number of specialist journals.

The mandate of a reviewer is inordinately difficult: to describe a book; to assess it; to inform generally; and perhaps to entertain. All this for love rather than money. It is galling for a reader and more so for an author of a book to find in several hundred words no mention of the work supposedly under review. In such an instance the reviewer has done a disservice to both parties. Yet the author should not complain if his book, as well as being judged implicitly or explicitly, has been used as a stage from which to enlighten, amuse or provoke. The reviewer may not have succeeded entirely in blending the functions of a review, but that is a measure of the difficulty of the task.

Oscar Wilde's approach to book reviewing was first to write the review, then, if he liked it, to read the book. That is not to be recommended; priority should go to assessment of the volume under consideration. But Wilde's technique does have a tinge of merit. □

Tim Lincoln is the Book Review Editor of Nature.

Brouhaha among the breadfruit

Paula Brown Glick

Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth.
By Derek Freeman.

Harvard University Press: 1983. Pp.379. \$20, £11.95.

FIFTY-five years after the publication of *Coming of Age in Samoa*, Derek Freeman has written this book to demonstrate that Margaret Mead's account of Samoan culture is "fundamentally in error", that she is incorrect in nearly everything she wrote about Samoa. The criticism is especially directed to *Coming of Age in Samoa*, but also refutes many of her statements in articles and in another of her books, *Social Organization of Manu'a*. Freeman's persuasive writing combines quotations from historical and ethnographic reports on Samoa with his own field research, documentary study in Samoa — including police and court records — and conversations with Samoans. He quotes anthropologists, psychologists and biologists on general issues, and makes extensive use of the writings of Margaret Mead, including her autobiography, *Blackberry Winter*, and *Letters from the Field*.

The book's publication was heralded by newspaper and magazine articles, radio and television appearances, as may be appropriate to Mead's standing and influence in the United States. In the years following her work in Samoa she conducted many other studies. She was a leader in techniques of observation, filming and recording behaviour. Her written work, as catalogued at New York's American Museum of Natural History where she worked throughout her life, runs to 1,597 items. Mead held a high position in American scholarship, received many awards and honours, was elected to distinguished scientific bodies. To her last days she made major speeches at scientific meetings. Throughout this time, she was in continuing demand as a public speaker, interviewed and quoted, wrote popular articles and magazine columns. She was the best known and most influential social scientist in America.

Freeman, now Professor Emeritus of Anthropology in the Research School of Pacific Studies at the Australian National University, has long been respected for his scholarship, particularly for his work on the Iban of Sarawak, who were the subjects

of a study he conducted in 1949 and 1950. His essay "On the Concept of the Kindred" won the Curl Bequest Prize of the Royal Anthropological Institute in 1960. A New Zealander, he conducted research in Samoa in the 1940s and made a series of later visits, the last in 1981. In the preface, he tells of a meeting with Mead in 1964, in



which he told her of his disagreement with her findings. He says that shortly before her death in 1978, he wrote, offering to send her a draft of his conclusions. His investigations have extended over 40 years, including six years in Samoa and even longer in research libraries. The documentation in his book is thorough: 55 pages of footnotes and many references to his own inquiries in Samoa.

Why should Freeman, a respected senior scholar, devote years of painstaking research to demonstrate that Mead's books and articles published over 50 years ago are false? There could be several answers, one of which is his championing of the

biological bases of human behaviour.

Freeman begins his book with a discussion of American anthropology in the 1920s. Franz Boas of Columbia University, then the leader of American cultural anthropology, was a strong advocate of cultural determinants of behaviour. Margaret Mead was his student. Early chapters describe an academic intrigue: Boas sent Mead into the field to prove his position in the nature-nurture controversy of the 1920s. Boas hoped to convince social and behavioural scientists of the time that culture is the dominant influence in personal development, with evidence from Samoa that adolescence need not represent a period of crisis or stress. In directing Mead's Samoan research Boas instructed her to concentrate upon adolescent girls, not to do a full ethnography of Samoa. She spent nine months in Tutuila and Manu'a, American Samoa. Immediately upon her return, Mead wrote *Coming of Age in Samoa*, which became an all-time bestseller, going through numerous editions and translations, and being assigned to countless undergraduates, quoted and expanded upon by other social and behavioural scientists; the book is still found in nearly every American library and bookstore.

To Freeman, Mead's Samoan research became the basis of an anthropological myth, which denies any role to biology or human nature in accounting for adolescent stress and human behaviour. In the 1920s Malinowski and Radcliffe-Brown, in Great Britain and the Commonwealth, were beginning to draw up the anthropological approach which has become known as "structural-functionalism". This holistic viewpoint aims to demonstrate the interconnections of activities and institutions within a social system. However, they had not yet had much impact in the United States. Culture, the particular beliefs and practices of people, remained of dominant interest in American anthropology. After the death of Boas in 1942, and of Ruth Benedict in 1949, Mead carried the ideas of cultural anthropology in her writings and speeches to an ever-wider public.

While Freeman's concluding chapter returns to the theme of biological influence on behaviour, the major part of the book details his objection to Mead's statements about Samoa. Repeatedly Freeman makes such statements as:

In the chapters that follow evidence will be adduced to show that the main conclusions of *Coming of Age in Samoa* are, in reality, the figments of an anthropological myth, which is

deeply at variance with the facts of Samoan ethnography and history [p.109].

However, the data sources and evidence presented by Mead and Freeman are quite different. Mead, then 23, worked with adolescent and pre-adolescent girls in three villages in Manu'a. In an appendix she presented data sheets for the analysis of girls' age, family and social position, and behaviour. She described adolescence as the best period of a woman's life, without stress or conflict, preoccupied with sex and lovemaking. Mead contrasted the gentle and carefree adolescence of Samoans with the educational and social pressures upon American youth, urging America to be less demanding of its young.

Freeman goes to great lengths to disprove Mead's characterization of Samoan adolescence. To establish that Samoan youths have a similar pattern of violence to that in other countries, he cites court records of Western Samoa, for 1967, which show that males in the age-group 14–20 constitute a high proportion of first offenders. It is as though observations of teenagers in a village such as Lower Slaughter in 1925 are disproved by Birmingham police records for 1967. Freeman says,

Mead, then, was at error in her depiction of the nature of adolescence in Samoa, just as she was . . . in her portrayal of other crucial aspects of Samoan life. This being so, her assertion in *Coming of Age in Samoa* of the absolute sovereignty of culture over biology, on the basis of these erroneous depictions, is clearly invalid, and her much-bruited "negative instance"* is seen to have been no negative instance at all. In other words, Mead's presentation of Samoa as proving the insignificance of biology in the etiology of adolescent behavior is revealed as a false case [p.268].

Freeman's critique includes the complaint that Mead never returned to Samoa, nor revised her books on the basis of later studies. In response to this, Mead speaks for herself in the preface to the 1973 edition of *Coming of Age in Samoa*:

Some young critics have even asked me when I am going to revise this book, and look unbelieving and angry when I say that to revise it would be impossible. It must remain, as all anthropological works must remain, exactly as it was written, true to what I saw in Samoa, and of what I was able to convey of what I saw. True to the state of our knowledge of human behavior as it was in the mid-1920s, true to our hopes and fears for the future of the world [p.x].

Further, Mead clearly acknowledged the importance of the work of others, including Freeman, in her conclusion to the reissued *Social Organization of Manu'a* in 1969:

Then, in 1965, I met Derek Freeman for the first time, and he challenged my material on the very mild Manu'an reactions I had reported on the

*The "negative instance" is defined by Freeman as Boas's view of scientific method in which a single negative case invalidates a hypothesis. In fact, Freeman dedicates his book to Karl Popper because of Popper's work on scientific method.

subject of virginity. He cited intense feeling about virginity on the part of mothers of girls, and extreme preoccupation with the theft of girls in other villages on the part of young men in western Samoa. In thinking this over, I realized for the first time that the whole of the rivalry between the young men or the chiefs of different villages, could be expressed in the hope of one group that they could do irreparable, *asymmetrical* damage to the other group. . . . Translated into terms of behavior, it accounts for the continuous wariness and touchiness which Samoans display in several contexts ever alert lest something one would enjoy doing to others be done to oneself. . . .

We need much more detailed material on early childhood, material which hopefully Derek Freeman's current fieldwork will provide [p.227].

The very nature of ethnographic field work, as a local study by a lone anthropologist, raises questions about reliability and generalization. The field worker pursues particular subjects and problems, chooses a community, works closely with only some members of that community, and publishes observations and conclusions based upon the fieldwork and the information available from other sources. Subsequent research among the same people — in this case an island group with a population of over 150,000 — is likely to find differences attributable to local conditions, historical changes, and interpretation or emphasis. Certainly ambiva-

lences and contradictions characterize every society. Conflicting findings are sometimes taken as evidence that anthropological research is both subjective and unreliable. And while Freeman presents much evidence to contradict Mead's statements, he cannot claim to have reproduced her conditions of field work with adolescent girls in Manu'a in 1925. His contention that she was innocently duped and misled by village girls is unwarranted.

Why, one wonders, has not Freeman written a new book on Samoa which incorporates the results of his researches, rather than a series of chapters devoted to Mead's "errors"? He amply demonstrates his ability to describe the social, political and rank system of Samoa in a way that summarizes his research and that of many other scholars into a clear and consistent statement. One can only surmise that his was personal preference for criticism over construction. By this choice, he has prevented anthropologists from evaluating the potential of his biological explanation of Samoan behaviour. Mead's work will continue to reflect its time and place, but Freeman has yet to demonstrate the correctness of his position. □

Paula Brown Glick is Professor of Anthropology at the State University of New York, Stony Brook. She is author of *Highland Peoples of New Guinea* (Cambridge University Press, 1978).

Little room for the anthropologist

Mary Douglas

Margaret Mead: A Voice for the Century.

By Robert Cassidy.

Universe Books: 1982. Pp.176. \$12.50.

MARGARET Mead, born into an academic family in Philadelphia in 1901, died internationally famous and loaded with honours in 1978. Her remarkable life deserves to be recorded and reflected upon. She was an anthropologist, a teacher, a humanitarian, a popular writer, a media star and a lay preacher. Her record defies classification but it invites a more reflective treatment than is here provided by Robert Cassidy. His book suffers the usual limitations of hagiography, strong in praise of the deceased and weak in dealing with the enigmas and problems of a person living in a particular place and time. We are left with hints — apparently fellow ethnologists accorded her only "grudging admiration" and her success was definitely in the larger community, "grandmother to America", "citizen of the world" or "voice for the century".

Her graduate work was in the University of Columbia, with Franz Boas and Ruth Benedict. Boas was building up a new

science of anthropology, specially focused on the relationship between the individual and culture and on methods of assessing that relationship. Cassidy's account of how Boas set the theoretical problems of her fieldwork project does not quite ring true. That towering scholar, geologist, physicist, organizer of great Arctic and north American expeditions, apparently first drew her attention to the cultural variation in sexual mores. He is said to have asked her whether adolescent troubles were peculiar to Western culture — how did young girls react to the constraints of custom, was it true that primitive girls were excessively bashful, did they develop crushes and other forms of romantic love? It sounds more like what his student told him she intended to study and clearly, even at that youthful age, she was not one to be easily diverted from her purposes.

In 1925 she went to Samoa, chosen by her supervisors as a place where she could be safe under the watchful eye of the United States Department of the Navy. Mary Slessor and Mary Kingsley, much earlier travellers in West Africa, and Edith Durham, travelling in the feuding Balkans, would not have made much of the dangers she faced in the South Pacific. On her return she wrote *Coming of Age in Samoa* published in 1928, which overnight became a bestseller. Perhaps the grudgingness of colleagues' admiration stemmed from this instant popular success, or perhaps it confirmed a wilful independence and sense

ADVERTISEMENT



Blue butterflies of the *Lycaenopsis*-group

By J.N. Eliot & A. Kawazoe
A detailed taxonomic study and revision.
Unique specimens and the rarer species
illustrated in colour. Line drawings, keys
and descriptions for identification to
subspecific level. 1983. 296pp. 6 plates
colour, 12b/w. 560 line figures.
0 565 00860 9 Hardback.
Until June 1. £25.00
Afterwards £28.00

A directory of natural history and related societies in Britain and Ireland

Compiled and edited by Audrey Meenan
Addresses, aims, amenities, meetings etc of
local, national, school and college societies
listed from A-Z, geographically and by
subject. "Natural History" being here
interpreted in its widest sense.
1983. 407pp. 0 565 00859 5
Hardback £15.00

A checklist of the helminth parasites of marsupials and monotremes

By E.A. Harris
Brings together the published literature on
the nematodes, trematodes, cestodes and
acanthocephalans recorded from marsupials
and monotremes. Microfiche, 2 fiche in
introductory folder. 0 565 00874 9 £6.00

List of British vascular plants

Prepared by J.E. Dandy
for the British Museum (Natural History)
and the Botanical Society of the British Isles.
Incorporating the London catalogue of
British plants. First printed 1958 this
important work has been out of print for
some years. It has now been reprinted to
meet continuing demand. v + 176pp.
0 565 00449 2 Hardback £5.00

Flora Zambesiaca.

Vol. VII part 1.
Families Crassulaceae, Valerianaceae,
Dipsacaceae; then Campanulaceae —
Loganiaceae.
Edited by E. Launert.
The latest published in this ongoing series by
the Managing Committee of Flora
Zambesiaca. 1983 394pp. 89 line drawings,
colour frontispiece. 095076 82 00. Paperback
£23.50.

Guernsey's earliest flora

By D. McClintock
An account of the plants of the Guernsey
countryside circa 1790. Based on the work of
Joshua Gosselin (1739-1812). 210pp,
frontispiece, 7 figures. A Ray Society
publication No.155. 0 903874 17 2
Hardback £14.00

Publications Sales, British Museum (Natural
History), Cromwell Road, London SW7
5BD.

of power in the young scholar. She was offered a post of Assistant Curator in the American Museum of Natural History in 1926 and declared that once she saw through the windows of her attic room in that august establishment she would never move out.

Was it really this resolution that prevented her from accepting any other academic appointment, as Cassidy implies? In 1954 she was made Adjunct Professor in the Anthropology Department at Columbia but never improved upon that part-time peripheral post until 1968, when Fordham University asked the now ageing superstar to develop their department of anthropology. There are others who have done great research from museums, and others who have been delayed from holding responsible office by the slow promotion in the museum world (as to academic tenure, Margaret Mead's rise was not meteoric; Associate Curator in 1942, Curator only in 1964). Others, again, prefer museum appointments for the freedom they give.

The interesting questions which Cassidy never attempts to broach are counterfactual conditionals about the interplay of sex, personality and culture, the very theme of Mead's life work. How much did she choose to be an outsider, how much was the choice a response to the opportunity structure she found herself facing — it is impossible to say. Yet, a man who showed such an early brilliance and energy would have been sought by major institutions which would have forced the conflict of ambition and independence to be played within supporting structures eminently able to bestow rich rewards on their favourites. Margaret Mead remained an individualist, she resisted capture by any institution and she exerted influence from outside the Establishment.

From her attic in New York Mead made a field trip of six months with her second husband, Reo Fortune, and published *Growing Up in New Guinea* in 1930. The path of research for the next ten years had been set. It was to show the malleability of human personality, the strength of cultural pressures to mould what are assumed in any one civilization to be personality traits indefeasibly linked to physical sex. The comparison of adolescence in Samoa and Manu'a was followed by comparison of three societies in New Guinea: the Arapesh, whose culture shows practically no differences in temperament between men and women, both being equally gentle and unaggressive; the Mundugumor, both sexes equally violent and ruthless; and the Tchambuli, a society where the men were artists and ran the ceremonies while the women were the real producers and entrepreneurial managers, the perfect complementary case to modern America.

One can pause to ask why she spent so little time in each of these societies. After all, this was the post-Malinowski era of long, intensive fieldwork. How, then, could

Cassidy not be sensitive to the import of his own remark: "After seven months she and Fortune had had their fill of the Arapesh". She felt she worked so much faster than anyone else; like a scientist, she chose a limited problem and could quickly tell whether she had collected enough evidence to settle the answer. Margaret Mead often complained that the British anthropologists were unappreciative of her achievements. This hurry-hurry, single-issue style of field work explains a significant difference between herself and her

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Margaret Mead in the late 1920s, after her return from Samoa.

British contemporaries — compare Audrey Richards, who wrote large monographs with impressive documentation and analysis on one tribe alone, and trained her students to keep going back and amplifying the results.

After her third marriage she went to two new field areas, Iatmul and Bali between 1936 and 1939. At this stage she had become interested in the techniques of ethnography, and the use of photography and film. In the 1930s she was one of the founders of the culture-personality movement in anthropology and was becoming very interested in national culture, but Kardiner and Dubois did the more important work in the first and Benedict in the second. She was there at the beginning of many intellectual movements but did not stay with them. One suspects that the demands of the mass media were beginning

to pose an effective counter-attraction to obscure scientific study.

By the outbreak of the war Mead was a very famous anthropologist. New challenges were put to her. As executive secretary for the National Research Council she edited the *Report of the Committee on Food Habits (1941-1943)*, full of interesting ideas about methods of comparing food habits and ideas about the relation between nutrition, social class and ethnic traditions. She advised the United Nations army on its relations with foreign cultures; she became an authority on technology and economic development and made it her prime mission to bring emerging Third World countries into the twentieth century. A peculiarly personal mantle of international counsellor and interpreter of the world fell upon her shoulders. No one wearing it with such authority and confidence should complain of lack of academic recognition. Yet I would pay her the compliment of greatly wishing that she had stayed more patiently with the growth of the ideas she brought to birth. Nothing in economic anthropology had anticipated her study of economic structures in the Admiralty Islands. She was right in believing at the end of her life that nothing significant had been built upon her important work on food habits. The outstanding problems about the relation of personality and culture are still unsolved and no method has been determined for solving them. But in academic life you either need to have institutional support or to concern yourself with the nurture of your own brain children.

It is fair to wonder with whom Mead should be compared. Was she a latter-day Marie Stopes? No, even their shared vilification as corruptors of morals does not put them on a par; Mead was less on-track. What about contemporaries? Was she a great philanthropic writer? Dorothy Day is a giant name for an editor and committed social reformer, worthy for the comparison, but Mead's concerns were more dispersed. What about helping the Third World? Margery Perham, born in 1895, achieved much on behalf of the constitution framers of emerging independent Africa? Certainly, I doubt that any of them would have envied Mead the international popular acclaim that she earned.

This biography pays scant attention to her anthropological work. It is clearly meant to be a friendly report. But it is not helpful to her academic reputation, now under attack by Derek Freeman (see p.758), to present her as a fabulous media star. Would Danny Kaye rather be remembered by his friends as a great comedian or as a great philanthropist? Margaret Mead herself cherished her academic standing, but this volume does nothing for such concerns. □

Mary Douglas is Avalon Professor of the Humanities at Northwestern University, Illinois.

Advancing by stealth

June Goodfield

Women Scientists in America: Struggles and Strategies to 1940.

By Margaret W. Rossiter.

Johns Hopkins University Press: 1983.

Pp.438. \$27.50, £22.

THERE are certain intriguing resemblances between the practice of science and the study of history: the more we observe, the more we recognize how much still remains to be discovered. Moreover what we actually see is strongly influenced by the questions we bring to our observations. Thus we can by-pass a great deal that may actually be very significant.

As directions in research become first fashionable, then hallowed, other aspects of a discipline may dwindle to near invisibility. It takes a Braudel, with his sweeping studies on the Mediterranean World, or a Ladurie in *Montaillou* and *Love, Death and Money in The Pays d'Oc*, to recreate lost lives and reveal the sheer dimension and complexity of the tapestry of history. Such was the majesty of their scholarship, and such the vision of their disciplined imagination, that the works of these pioneers forced those with minds not irrevocably closed to look with fresh eyes on old fields of study.

While it cannot be said that Margaret Rossiter is a Braudel or a Ladurie, her new book nevertheless achieves similar ends and is a seminal work of rich scholarly detail. It may not transform the history of science, but those who ignore its implications will be intellectually impoverished. Not too long ago many people believed that Marie Curie was the first — if not the only — woman scientist. The sixteen-volume *Dictionary of Scientific Biography*, edited by Charles Gillispie, did something to improve the situation but nothing like enough. Dr Marilyn Ogilvie's labour of love — lasting over a decade — which has culminated in a *Dictionary and Biography of Women Scientists from Antiquity to 1910* will, when it finally sees the light of publication day, open our eyes still further. Now with *Women Scientists in America* Margaret Rossiter reveals that for well over a century "women have been an integral part of the scientific community" yet "for a variety of reasons" most of these women scientists bordered on the "invisible".

Once she had overcome the initial "stumbling block of locating material, through several years of detective work" the full wealth of the vein she was mining became apparent. The quantity of material discovered was enormous; once mastered, rich and deep analyses could be and were, undertaken. Thus Rossiter's account is not one whose only conceptual threads are those of feminist rhetoric. There is nothing simplistic here: she has given us a sophisticated analysis of a complicated situation,

encapsulated in the following phrase: "It is the history of an occupational group whose status had risen and fallen over time as the women's role responded to external events and pressures". What Margaret Rossiter does so well is to explicate her story in those terms, backed by an abundance of empirical material, in quantity that would satisfy the most Baconian of scientists.

Beginning in the nineteenth century with the "entering wedge" provided by the women's colleges, we see the quiet infiltration of women into science, a movement that turned out to be unexpected, unintended and at certain times unwelcome. It is fascinating to notice that, at a time when science had a lowly status in Britain's institutes of learning, the Women's Colleges of America took pains to emphasize the extent of science in their curricula offerings, showing a slant undreamt of by C.P. Snow. For the emphasis placed on science "reassured the parents of the high moral character and religious values taught. (Literature, especially novels, were, by contrast, quite suspect.)". Given this difference in educational history, surely no one would ever have predicted that Britain would produce the first scientist ever to be a Prime Minister and a woman at that!

The ends which were ultimately achieved were far from being those for which the new ventures were launched. What began as a movement to educate women so that they would, in turn, produce more enlightened sons, ended with a group of highly motivated, qualified and often quite remarkable people having no place to go except back to where they came from — the women's colleges. This, rather than any inherent intellectual limitation, explains why few of these women scientists did research — though there were notable exceptions. That this situation, and its corollary, persisted for well over one hundred years is confirmed in Evelyn Keller's forthcoming biography of Barbara McClintock. McClintock's work in bacterial genetics was so brilliant that Joshua Lederberg, in the 1950s, once described her as either "mad or a genius", but she was forced to do her research on the fringes of academia. For she refused to relinquish her calling, even though appropriate positions in academic institutions were never forthcoming; nor would she abandon her lodestar and fall back into teaching science in a women's college.

Rossiter clearly displays the causes and

Gentlemen of Science

Oxford University Press have recently issued a paperback edition of Jack Morrell and Arnold Thackray's *Gentlemen of Science: Early Years of the British Association for the Advancement of Science*. In his review of the original edition (*Nature* 294, 19; 1981) Robert Fox spoke of the book as bristling "with challenging interpretations". Price of the paperback is £7.95.

effects of the fluctuations in women's status in the American scientific scene, as well as the social and personal obstacles that lay in their paths. Many men, such as Amos Rennselaer, helped from deep conviction; many women, such as Dean Laura Gill of Columbia, hindered — equally from deep convictions. With humour and barbed shaft, Margaret Rossiter has provided a book that, in fact or analysis, in reference or bibliography, is a splendid and totally satisfying feast, whetting the appetite for the next volume that will bring the record from 1940 up to the present day.

Beyond the present who can say what new pressures may influence the status of women in science, whether in Europe or in America? There are some who argue that the position is now getting worse again as the pendulum swings once more. Rossiter's wry comment on the early pioneers was true in the nineteenth century and in the prime of Barbara McClintock's life in the 1930s: "If one wishes to introduce seemingly radical reforms it helps to be personally discreet and socially conservative". There is no reason to suppose that it is not true now. *Plus ça change...* □

June Goodfield is Visiting Associate Professor of Biomedical Science and Social Policy in Medicine at Cornell University Medical College. Among other books she is author of *Reflections on Science and the Media (AAAS, 1981)*.

ADVERTISEMENT

TELL THEM HOW IT REALLY IS

"THE ATOM: A CLOSER LOOK"

**A Unique Film
About Atomic Energy**
30 minute 16mm color film
to introduce students, grades
6 through 12, and the general
public to nuclear energy.

Describes atom's structure, nuclear
chain reaction, the operation of
nuclear power plants, radiation's af-
fect on the environment and nuclear-
waste disposal.

PRODUCED

BY: **WALT DISNEY**

Educational Media Company
English German
Neutral Spanish

American Nuclear Society

555 North Kensington Avenue 312/352-6611
La Grange Park, IL 60525 USA Telex: 254635

Place for a quiet revolution

Derek A. Roe

The Middle Stone Age at Klasies River Mouth in South Africa.

By Ronald Singer and John Wymer.
University of Chicago Press: 1982.
Pp.231. \$30, £24.

FROM time to time, discoveries are made which radically change our perception of some segment of the past. It happened, for example, in the 1960s, when potassium-argon dating suddenly almost doubled the Lower Palaeolithic time-scale and the estimated age of the Plio-Pleistocene boundary. Over the past 15 years, there has been an almost as drastic if rather more gradual reappraisal of the age and status of the sub-Saharan Middle Stone Age (MSA). Klasies River Mouth, on the Tzitzikama coast of South Africa, is one of the crucial sites in this quiet revolution, and this volume gives a definitive account of the archaeological, geological and environmental research carried out at the group of caves and rock-shelters there. Hitherto, only the fauna and sediment stratigraphy had been reported on, so this publication had been eagerly awaited.

The conventional view of the MSA of southern Africa, virtually unchallenged even ten years ago, made it broadly contemporary with the Upper Palaeolithic of the Northern Hemisphere, beginning perhaps in the 30,000s and possibly lasting until the end of the Pleistocene. True, the stone tools did mostly resemble those of the much earlier Middle Palaeolithic in the north, where many people would have called them "Levalloisian" or "Mousterian"

without a second thought, but it was assumed that in southernmost Africa there had been no "real" Upper Palaeolithic and old techniques would therefore have lasted on. The extent of the reappraisal is that we now see the MSA as well established by 120,000 years ago, perhaps substantially earlier; in most places it must have ended well before the date when it was previously assumed to begin. It has therefore become a chronological as well as a technological equivalent of the northern Middle Palaeolithic, which makes excellent sense. The South African Late Stone Age, incidentally, has expanded to fill most of the gap, but that is another story.

The role of Klasies in all this has been to provide by far the longest and richest MSA sequence, with its earliest stage closely related to a marine transgression of "Last Interglacial" age, almost certainly of oxygen isotope stage 5e (120–130,000 years ago). The caves and shelters are interpreted in combination to yield a remarkable single sequence, including an unbroken succession of deposits no less than 18 metres deep, with signs that a further 6 metres may have been lost by erosion. More MSA deposits follow, before a long break, which may exceed 50,000 years; afterwards, "strandloper" middens accumulated, rather late in the Late Stone Age. The Middle Stone Age sequence is divided into MSA I–IV, the four stages showing only minor variations in their tool-kits, which were based on the production of flakes and flake-blades from prepared cores. The erosional break referred to above comes between MSA III and IV. But between MSA II and III comes a substantial and important occurrence of what is called Howieson's Poort material — a strikingly different and forward-looking industry, with smaller and narrower flake-blades, probably punch-struck, many of them worked into

Years BP	Environment	Archaeology
c. 2,500	Sea level near present. Dry climate.	LSA II shell midden
c. 4,700	Sea level near present. Cave damp.	LSA I shell midden
c. 5,000 	Major stalagmite formation. Erosion of external midden and front of cave deposits. Sea near present level.	No occupation
c. 65,000 c. 70,000	Major drop in sea level: oxygen isotope stage 4. Wet. Aeolian sand covers the cave deposits.	Sparse occupation of MSA IV
c. 80,000	Dry	MSA III
c. 95,000	Sea near present level. Cooler and dry.	Howieson's Poort industry
c. 105,000	Sea level at 4–5 m O.D.	MSA II
c. 115,000	Sea near present level.	MSA II
c. 120,000	Slight fall in sea level.	Initial occupation of MSA I, with <i>Homo sapiens sapiens</i>
c. 125,000	Main cave eroded or re-modelled at time of major transgression of 6–8 m sea (oxygen isotope substage 5e).	No occupation

Klasies River Mouth, Eastern Cape Province, South Africa. Time scale based on relative sea levels, environmental interpretation and external stratigraphic inferences by Karl W. Butzer, correlated against major archaeological divisions by John Wymer. The Middle Stone Age (MSA) of South Africa is based on production of substantial flake-blades from prepared cores, resembling the Levallois technique as found in Europe, the Levant and North Africa. The Howieson's Poort is a semi-microlithic industry with small scrapers, graters and various backed blade types. It is the earliest known industry of this kind, with certain technological similarities to the Upper Palaeolithic industries of Europe and western Asia which go back to c. 35–40,000 years ago.

"semi-microlithic" tools such as crescents and trapezes. The stratigraphic position of this material is unequivocal. Discussion of the dating and geological correlations of the Klasies sequence allows for "long" and "short" chronologies at this stage; the shorter reading would make the Howieson's Poort industry 30–50,000 years old, but the authors prefer a longer chronology, which would entail an even more surprising age of c. 95,000. When work at Klasies River Mouth began, the few Howieson's Poort occurrences known were regarded as transitional between the Middle and Late Stone Ages.

In spite of Upper Pleistocene marine fluctuations, the site remained on or near the coast throughout the MSA; its long abandonment after MSA IV probably corresponds to the main regression of the Last Glaciation. The MSA levels document in detail the evidently highly successful

results achieved at Klasies River Mouth, which may indicate the importance of this publication. Excavation at Klasies began in 1966 and marks, with the work of certain others at about that time, the start of a new and extremely fruitful period in South African archaeology, which has since seen many excavations of high quality, backed by painstaking and imaginative laboratory work of many kinds. A tremendous amount of new information has been won, with consequent radical changes in our understanding of the local sequence over

almost the whole prehistoric period. The surge forward is continuing and there will be much more to come even during the present decade. Meanwhile, Ronald Singer, John Wymer and their colleagues are to be congratulated on their achievements at Klasies River Mouth, so clearly set out in this handsome and well-illustrated volume. □

Derek Roe is University Lecturer in Prehistoric Archaeology and Director of the Donald Baden-Powell Quaternary Research Centre at Oxford University.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Changing our perception of the past — the Klasies River Mouth cave and shelter complex.

human exploitation of coastal resources, for which there is little evidence anywhere from earlier sites. Collecting of shellfish predominated; marine mammals and sea birds were hunted and stranded fish taken, but there seems to have been no regular fishing.

Various hominid remains also came from the MSA levels: no burials, alas, but quite a worthwhile aggregate of bits and pieces. The authors are satisfied not merely that *Homo sapiens sapiens* was established in this area by the end of the MSA, but also that he made the earliest MSA industry at the site. If so, we presumably see here the earliest representatives of that sub-species yet discovered; their presence in southernmost Africa so early will clearly need to be taken into account in all theories of the final stages of human evolution further north. Sadly, no hominid remains occurred with the Howieson's Poort industries, which present a classic interpretive dilemma: do the innovative changes in the tool-kits represent indigenous adaptations to new ecological circumstances, or do they imply a new and intrusive population? The authors clearly prefer the latter interpretation, and associated hominid material might have offered decisive evidence.

These, then, are some of the far-reaching

Shaking behaviour

Chi-Yu King

When the Snakes Awake: Animals and Earthquake Prediction.

By Helmut Tributsch.

MIT Press: 1982. Pp.248. \$27, £18.

REPORTS of strange behaviour by animals shortly before earthquakes are common. For example, cats picked up kittens and left houses; dogs barked madly; pigs bit each other; fish jumped out of water; and snakes awoke from hibernation, crawled from their burrows and, in some cases, froze to death on snow-covered surfaces.

Until recent years such reports and suggestions that abnormal behaviour might be related to earthquakes were largely ignored by Western scientists, primarily because most of the observations were anecdotal in nature and made by people who were not scientifically trained and might have been biased by the excitement after the earthquakes. Besides, it was not clear how animals could sense the coming of an earthquake, while human beings and their instruments commonly could not.

However, as more and more observations of this kind were made known, a small but growing number of scientists began to examine the possibility of animals' earthquake premonition. In 1976 and 1979, the US Geological Survey, the Government agency in charge of the national earthquake-hazard-reduction programme in the United States, convened two separate scientific conferences on this topic. In China, abnormal animal behaviour not only was observed by numerous people before earthquakes but also has actually been used as one of the very few methods to predict earthquakes on a short-term (hours to days) basis since the 1966 Xingtai earthquake.

One of the scientists who have been recently intrigued by this phenomenon is Helmut Tributsch, a physical chemist, who learned of abnormal animal behaviour from the survivors of the 1976 Friuli earthquake that destroyed his home village in northern Italy. He then diligently searched the literature and collected anecdotal reports on 77 other earthquakes. His findings are presented in *When the Snakes*

Awake, originally published in German in 1978. In the book he also describes observations of several other phenomena thought to precede earthquakes — fog, lights and sounds — and discusses various hypotheses. On the basis of the available circumstantial evidence, he makes a strong argument for the attribution of these phenomena to an increase in the number of electrostatically charged particles in the atmosphere, generated by the tectonically stressed crust.

Written for laymen as well as the scientist, the book contains some simplified geophysical statements that may not be rigorous enough for the experts. Seismologists, for example, will probably be annoyed by Tributsch's frequent use of the words "force", "intensity" and "strength" interchangeably with "magnitude" in describing the size of earthquakes; volcanologists may disagree with the statement that "volcanic eruptions are special cases of earthquake activity" (p.55); and physicists will be dissatisfied by the absence of direct evidence for the postulated release of charged aerosols from the crust and by the absence of a credible mechanism to account for such a process. In addition, more consideration should probably have been given to the possibility that animals react to several geophysical and geochemical stimuli, and of a tectonically induced increase in the outgassing rate before earthquakes. Production in the biosphere of a significant amount of deep-earth gas, which differs considerably in chemical composition from the normal atmosphere — less oxygen and more methane, water vapour and radon — seems to be quite capable of causing not only most of the above-mentioned pre-earthquake phenomena but in addition the postulated charged-aerosol increase itself.

Irrespective of the above-mentioned criticisms, this book is, to my knowledge, the most comprehensive treatment on the subject in English; Tributsch's writing is rational, lucid and interesting. I expect that the book will stimulate further studies by scientists, as well as catch the attention of the general public living in earthquake-prone countries. □

Chi-Yu King is at the US Geological Survey, Menlo Park, California.

Geology, but not on a plate

Peter J. Smith

In Suspect Terrain.

By John McPhee.

Farrar, Straus & Giroux: 1983. Pp.209.
\$12.95.

JOHN McPhee is a staff writer for *The New Yorker* who specializes in writing book-length articles that describe and interpret to the world in general the activities of particular individuals and groups within it. In 1980, for his fifteenth investigation, he turned his attention to geologists, particularly those concerned with the Basin and Range province of the United States, to see how the earth sciences in their post-revolutionary phase differ from the traditional sort of geology that he himself had been taught in an introductory course several decades ago. The result, *Basin and Range* — first published by Farrar, Straus & Giroux, and earlier this year in Britain by Faber — was so successful that it led to McPhee's being feted by the North American earth-science community, though perhaps more for his uncritical, awe-struck reverence than for his ability to capture the spirit of modern geology on the page.

For what the flattered earth scientists had failed to notice was that by concentrating largely on field geology, rather than portraying today's subject as a complex interaction of fieldwork, laboratory experiment and theoretical insight, McPhee was guilty of severely understating the most important of the cross-revolutionary differences he purported to be seeking. The view of geology and geologists he offered to the outside world was thus hardly undistorted. Moreover, for British tastes, his excesses of style were hard to take, not least his tendency to generate excitement artificially by gee-whiz prose with an abundance of one-word sentences and decidedly purple passages ('salt-and-peppery charcoal-tweed savings-bank rock'). Nevertheless, the book had undoubted strengths. McPhee's set pieces on the Huttonian and (more briefly) the plate tectonic revolutions, vivid expositions aided by a nice line in extended metaphor, were beautifully executed examples of precisely what one looks for from an external interpreter with literary aspirations.

With *In Suspect Terrain*, his second foray into the geological world, McPhee has now confounded my worst fears by producing a far superior work of remarkable power. The tendency to overblow is here much muted and the sober strengths of the previous volume are given greater play. Largely gone, too, is McPhee's irritating habit of using geological words and concepts without explanation, thereby

creating an impression of the author as an insider playing a game of one-upmanship with his non-geologist readers. The emphasis on field geology is still there, but now it has decided point — to extol the virtues of knowledge gained in the field as a counterbalance to the worst excesses of the gospel-thumping plate movers. Indeed, the geologist McPhee chose to guide him in the field, Anita Harris of the US Geological Survey, knows only too well the danger of being mesmerized by the plate people. In 1972 the pressures of the consensus led her and her colleagues to misinterpret the conodont palaeontology of an area of Pennsylvania so seriously that they came to see plate interactions where the evidence hardly warranted it — an episode that later led to painful withdrawal symptoms.

What Harris thought she had found in Pennsylvania was an exotic block, a piece of crust alien to its surroundings, a suspect terrain. Not that all suspect terrains are quite so suspect. The glacially-emplaced exotic materials with which McPhee begins and ends his spiritual journey are hardly likely to require major reinterpretation within any conceivable future. But at the

opposite extreme, the current passion for fragment tectonics and geosuturing, against which Harris so passionately inveighs in the context of the Appalachians, is altogether more insecure. At the heart of McPhee's exposition, therefore, is an extended evocation of the complex geology of the Appalachians — not a geological textbook or guidebook, but a brilliant impressionistic picture of Appalachian geology as seen on the ground acting as counterpoint to Appalachian geology perceived by the office-based global tectonicist.

"Plate tectonics is . . . a cop-out. It's what you do when you don't want to think", McPhee quotes Harris as saying. In highlighting this growing gulf between field and office geologists, McPhee is perhaps imparting to the general reader more of the spirit of modern earth sciences than some practitioners would wish. Be that as it may, he has produced by art a minor masterpiece of scientific communication. □

Peter J. Smith is Reader in Earth Sciences at the Open University, and editor of *Open Earth*.

Art of indoctrination

Glynn Isaac

The Creative Explosion: An Inquiry into the Origins of Art and Religion.

By John E. Pfeiffer.

Harper & Row: 1982. Pp.270. \$28.80, £18.50.

IN MODERN human beings the brain is the organ of culture. The intricacies of technology, communication and social interdependence would be impossible without the enlarged, elaborated brain of *Homo sapiens*; conversely having such a slowly grown, metabolically expensive organ would make no evolutionary sense in the absence of the complex programmes of rules and information, which make the human way of life possible. These programmes (including language) must be learned by each individual as he or she grows up. They comprise what anthropologists call culture.

Archaeology and hominid fossil skulls together provide evidence that, for the past two or three million years, the brain and culture developed as a joint system in a way that is partly analogous to the development of computer hardware and software over the past three decades. Looking at this long record as a whole, one can perhaps usefully pick out three, or maybe four, particularly significant transitions. First, two million years ago or earlier, tool-making, meat-eating and maybe some aspects of a basic human-like socio-economic pattern became evident; then for a very long time the record implies only very slow change: consolidation and some elaboration of these elements, but no fundamental shifts

(unless the acquisition of control over fire is to be so classed). The second transition occurred inside the past 100,000 years and seems to involve the emergence both of anatomically "modern" human beings and of levels of social-cultural complexity such as are familiar in history and ethnography. In as far as people know about this transition at all, it is associated with the passing of the Neanderthals and the Cro Magnon forms with their Upper Palaeolithic culture. The third and fourth transitions follow hard on the heels of the second; they involve food production (farming) and the development of expanded, intensified networks of organized interdependence that we call "civilization". In a sense the third and fourth developments can be seen simply as consequences of the surge of change and diversification that followed upon the second transition.

That second transition is the least well understood, and until recently was not even subject to clear-minded attempts to formulate questions. It is in this context that John Pfeiffer's book makes its contribution.

One of the most intriguing manifestations of the Upper Pleistocene transition is the first appearance of Palaeolithic art in

Solid state physics

In 1981 Addison-Wesley published a single-volume *Encyclopedia of Physics*, reviewed in *Nature* 290, 657 (1981). An offshoot of the parent volume which has just been published is the *Concise Encyclopedia of Solid State Physics*, also edited by Rita G. Lerner and George L. Trigg. The new book aims "to provide . . . a convenient and readily accessible set of resource materials" on the subject, and is available in hardback and paperback. Prices are £29.60, \$39.50, and £17.20, \$22.95 respectively.

the archaeological record. Twenty to thirty thousand years ago the oldest known, definitive instances of such art pop up in archaeological sequences in several far-flung parts of the world, for instance South Africa, Australia and Europe.

John Pfeiffer's book deals with two questions. Why art? And why at this point in the development of the human condition? He tackles them by considering one class of art in particular, namely the representations which were painted, engraved and sculptured deep in the caves of France and Spain. He considers that the appearance of the art can best be understood in relation to new and crucial social needs. More specifically, he argues that by 30,000 years ago the volume of information to be transmitted from generation to generation in some societies had risen to levels such that indoctrination involving the dramatic presentation of materials was essential to ensure accurate uptake and further propagation. He argues that the representations in caves are surviving evidence of rituals which not only helped to establish the authority of those who controlled the elaborate rules but also encoded parts of the programmes in unforgettable ways.

The view that the deep-cave art of Europe was surrounded by magico-religious ceremonies is not at all new. What is new is the explicit argument that during the Upper Pleistocene the information load to be transmitted had become so complex that special encoding and dramatic mnemonic systems become adaptive for the first time. John Pfeiffer is not alone in starting to think about such questions, his work joining the small but significant corpus of such scholars as Meg Conkey, Les Freeman, Clive Gamble, Michel Lorblanchet, Peter Ucko and Martin Wobst, among others. All of these have begun to be concerned with the problems of information flow in band societies, and with how archaeologists might monitor changing patterns in the Pleistocene. Pfeiffer's focus on the deep-cave art, however, and his argument for drama and situation as memory-imparting devices are, as far as I am aware, quite distinctive.

The book has several aspects: first it provides a brief synopsis of prehistory in general; second it presents vivid accounts of the author's own impressions and experiences in visiting the sites of deep-cave

art; third it describes in qualitative terms features of the art, its context and other, associated traces; and fourth it advances the arguments already alluded to regarding the novel social needs of Upper Palaeolithic societies. The book is clearly intended to be read by a general as well as a scholarly audience, and the first three features have evidently been included to make the book accessible to non-archaeologists. Expositions of the drama of the setting of the art are repeated in several parts of the book. I found this material convincing the first time but thereafter it may be that some readers, like me, will find it overdone and repetitious.

The book includes a chapter on myth, ritual and art in Australia, considered as an effective way of imprinting complex topographic, economic and social information.

Pfeiffer coins the apt phrase "tribal encyclopedia" for the sum of shared information learned and transmitted in non-literate societies.

As a non-specialist in cave art I found that the book included a variety of information that was new to me, for instance the fact that footprints of adolescents are commonly found in various deep caves where tracks are preserved. Such evidence adds plausibility to Pfeiffer's argument that initiation and indoctrination were important social functions of the art. Initiation is in fact a possible usage that had earlier been suggested by

Breuil, though without explicit consideration of the overall function of the transmission of information. Clearly key points of evidence such as the proportion of juvenile prints and other quantitative data will need to be more systematically presented by Pfeiffer, or some one else, before their implications can be fully assessed.

As indicated at the outset, the transformation of cultural systems during the upper Pleistocene seems to have been the launching pad for technological, economic, demographic and social changes that began to accelerate then, and which are still gathering speed. Pfeiffer is very probably correct in suggesting that new mechanisms for channelling the flow of complex information played a crucial part in this take-off, and he has produced a well-illustrated and provocative book about it.

Glynn Isaac is Professor of Anthropology at the University of California, Berkeley.

Reversals of fortune

F.J. Vine

The Road to Jaramillo: Critical Years of the Revolution in Earth Science.

By William Glen.

Stanford University Press: 1982. Pp.459. \$37.50.

THE MAIN title of William Glen's book is rather esoteric and perhaps a little deceptive in that it suggests a racy novel; the subtitle, on the other hand, is considerably more informative. Readers of fiction will not be entirely disappointed however for the story Glen has to tell, while firmly based in fact and soberly recounted, is a gripping one.

In essence the book covers events in three sub-disciplines of the earth sciences — geochronology, palaeomagnetism and marine geophysics — during the years 1954 to 1966; events which resulted in the definition of a geomagnetic reversal time-scale and the confirmation of the hypothesis of sea-floor spreading. Glen considers that these were the crucial developments which led rapidly to the general acceptance of the theory of continental drift and to the formulation of the modern paradigm of plate tectonics.

The Road to Jaramillo is therefore in the realm of the history and philosophy of science. It is, however, rather more than a straightforward history, based on months of work in the library, in that it draws heavily on the personal recollections, correspondence and records of the participating scientists. Thus the book provides a detailed and fascinating account not only of the way in which science proceeds but also of the way in which scientists behave while doing science.

In the 1970s several books were written on the Kuhnian type of "revolution" which took place in the earth sciences in the preceding decade. They included a book by Glen himself and were, in their different ways, faithful accounts. All of them, however, were broad-brush treatments, either designed to be of general interest or specifically aimed at undergraduates. Although the treatment in the latter part of *The Road to Jaramillo* is similar to that of such earlier works, the first two parts, which constitute more than three-quarters of the main text, are very different. They cover the refinement of the potassium 40-argon 40 dating technique between 1954 and 1960, which facilitated the dating of much younger rocks than had been possible previously, and the subsequent application of this technique, together with palaeomagnetic measurements, to geologically young lava flows in order to determine their age and the direction of their "fossil" or permanent magnetization, i.e. whether it was essentially parallel (normal) or antiparallel (reverse) to the present direction of the Earth's magnetic field. In



Femme à la Corne, a rock carving ascribed to the Upper Perigordian c. 23,000 years BP. The figure is 44 cm high and found in Laussel, a rock-shelter in the Dordogne. Photograph: B. Biraben.

this way a time table of reversals of the field during the past four million years was painstakingly built up, the last important detail — the resolution of the Jaramillo event, a short period of normal polarity rather less than one million years ago — being added in 1966.

It is made clear in the author's preface that the book is an outgrowth of a doctoral dissertation, on the history of the elucidation of the reversal time-scale, which was then expanded to include details of the all-important precursor to that work. As a geologist turned historian, Glen is well-qualified to relate this story and he is also well-placed in that these developments largely took place at institutions in the Bay area of California where he and most of the participants still live. Thus the author

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

F.J. Vine and Drummond Matthews at UEA in 1970. Their explanation, using the reversal time scale, of how the sea-floor developed a striped magnetic pattern (a theory developed independently by Lawrence Morley in Canada) eventually triggered the modern revolution in earth science.

was able to draw in particular on the recollections of Jack Evernden and Garniss Curtis — who refined the potassium-argon dating technique at Berkeley in the 1950s — Allan Cox, Richard Doell and Brent Dalrymple — all former Berkeley students, who worked on the definition of the geomagnetic reversal time-scale at the US Geological Survey in Menlo Park, California, in the early 1960s — and their more senior mentors at Berkeley, John Verhoogen and Francis Turner. Many others however were involved not only from the United States but also, notably, from Canada, Australia and Britain. Glen claims to have interviewed over 80 scientists for a total of more than 300 hours, 130 of these being preserved for posterity on tape. In doing this he has performed a valuable service to both scientists and historians; the book captures something of the sociology as well as the history of events before the record has been tidied-up, abbreviated and over-simplified by the passage of time.

Inevitably perhaps, because the book goes into a few topics in great detail, the general reader may find some of Glen's scientific and historical points somewhat

esoteric. One might also quibble with the odd minor detail or ambiguity, but in general the treatment is objective and factually correct, the whole constituting a scholarly and authoritative work. As one who was privileged to be involved with the latter part of the work described, I was fascinated by some of the details given earlier in the book which previously I knew only in outline. This is particularly true of the developments in the potassium-argon dating technique in the mid to late 1950s and may, I suspect, be a more general reaction in that the history and significance of these developments has been poorly documented, indeed largely overlooked, in the past. Similarly the sudden growth in palaeomagnetic research, and the eventual divergence in the approach of the English as opposed to the American/Australian schools in the 1950s makes interesting reading. Work in England under the leadership of Keith Runcorn concentrated on the definition of polar wander curves in the past for different continents; by contrast research in the States under Cox and Doell, and in Australia under the direction of Ian McDougall, was largely concerned with the elucidation of the reversal time-scale which was eventually to have such a dramatic impact on the resolution of the continental drift debate.

One is tempted to make comparisons with James Watson's *The Double Helix*, in some respects an analogous book. I suspect, however, that the success of the latter in capturing the popular imagination resides in its autobiographical approach and in its concentration on feuds and personalities. Here one can read about the rivalries, disagreements and personality clashes but they are more deeply buried in objectivity and detail. Indeed this is altogether a more weighty tome, painstakingly annotated, referenced and indexed; at the same time it should be said that the main text is attractively written and is leavened — and enlivened — by numerous photographs, typically of the leading characters in the story.

The road to Jaramillo was far from straight and well signposted and many travellers clearly missed crucial turnings on the way. But they all shared an enthusiasm for the unfettered pursuit of pure science, and had an instinctive faith in the ultimate, if unforeseen, fruits of their work. In this regard a salutary and poignant note is struck in an appendix in which Brent Dalrymple, in a letter to the author, questions whether such triumphs of pure science are as likely or even possible today in an atmosphere of financial stringency and in which the emphasis is on mission-orientated and applied research. □

F. J. Vine is a Professor in the School of Environmental Sciences at the University of East Anglia, Norwich. In the 1960s he was involved on both sides of the Atlantic in work which led to the confirmation of sea-floor spreading and the development of the theory of plate tectonics.

Wrendered complete

Carole Stott

The Mathematical Science of Christopher Wren.

By J.A. Bennett.

Cambridge University Press: 1983.

Pp.150. £15, \$29.95.

ASK THE man in the street to name an English architect and the chances are he'll say Christopher Wren. Follow this by asking for an English astronomer and your money is safe that he'll "pass" rather than answer Wren.

It is understandable why Wren is popularly acclaimed for his architecture. His most obvious legacy, his buildings — St Paul's, numerous churches, the Sheldonian Theatre, Oxford, the chapel at Pembroke College, Cambridge — are there for all to see. Yet before becoming a professional architect Wren had a highly successful career as an astronomer, firstly as Gresham Professor of Astronomy (1657-1661) in London and then as Savilian Professor of Astronomy at Oxford (1661-1673). These two seemingly distinct careers, architect and astronomer, have always been treated separately by historians. *The Mathematical Science of Christopher Wren* by contrast covers Wren's professional life as a whole, explaining what today would be a most unusual transition rather than, in those days, one of job rather than of career.

Whilst at Gresham College, Wren was at the historic centre of mathematical science in England, steeped in tradition concerned with the exploitation of mathematics to practical ends. The philosophy of the mathematical practitioners of the time and Wren's involvement are treated first by Bennett, before covering in successive chapters the mathematical sciences tackled by Wren — astronomy, longitude, cosmology, mechanics, microscopy, surveying, medicine, meteorology and architecture.

An early problem that concerned Wren was that of explaining the different appearances of the planet Saturn. Since Galileo had turned his telescope to the heavens and viewed what he thought was a central globe flanked by two smaller ones, astronomers had put forward various explanations for their observations. Wren observed the planet too and hypothesized that Saturn was circled by an elliptical ring the width of which varies from a maximum at the two points where it is furthest from the body to zero at two points where it touches the planet. Although this was a considerable advance on any previous explanatory attempt, Wren soon abandoned his theory in favour of Christian Huygens' model.

Self-devised improvements in the telescope helped Wren in his astronomical

work, including that on Saturn. He also used micrometers in the mid-1650s, he proposed lens grinding apparatus, the application of telescopic sights and the use of a double-telescope. The latter he described as two telescopes joined like a sector to measure angular distances and to be used in an astronomical method for finding longitude at sea. In all Wren was the complete scientist and his other interests — such as the anatomy of the brain, muscular action, respiration and designs for weather clocks — are all meticulously detailed by Bennett.

This is no light-weight tome, being the distillation of a doctoral thesis in the Department of the History and Philosophy of Science at Cambridge University. Some 20 pages are taken up by references, underlining the fact that it is aimed at the

professional scientific historian. The reference section is excellent and alone is ample reason for buying the book. The overall quality of production is up to usual Cambridge University standards and illustrations are good considering the limited availability of material. The author's assumption that his readers are familiar with scientific problems and personalities of the day, however, means that we are presented with a skeleton of a book. The present format unfortunately has an effect similar to force feeding; it is very rich food indeed and is only palatable in small quantities. One hopes that Bennett will in due course find time to fatten up his account. □

Carole Stott is Curator of Astronomy at the Old Royal Observatory, Greenwich, part of the National Maritime Museum.

Science through the scientist

Nicholas Wade

Scientific Temperaments: Three Lives in Contemporary Science.

By Philip J. Hilts.

Simon & Schuster: 1983. Pp.302. \$15.95.

PHILIP Hilts writes on science and kindred matters for the *Washington Post*. In *Scientific Temperaments* he has applied the reporter's art to describing the work and character of three contemporary practitioners of science. The result is a rich tapestry in which the substance and oral history of science are skilfully interwoven with the personal lives of individuals.

The book is composed of three essays, the first devoted to modern physics and Robert Wilson, the second to molecular biology and Mark Ptashne, and the third to computer science and John McCarthy. Each essay is based on a wealth of gracefully sketched background material, against which the subject's own career and outlook is presented. Perhaps because of the comparative youth of computer science, McCarthy is the member of the trio who stands out most clearly and strongly against his background.

Lord Rutherford, as is well known, pronounced in 1933 that the idea of nuclear power was moonshine. Less well known, at least to me, is that this was the year in which H.G. Wells had predicted (in 1914) that the means to tap the energy of the nucleus would be found. The prediction, as Hilts tells the story, deeply impressed Leo Szilard, but not until Szilard happened to attend the conference at which Rutherford declared nuclear power to be impossible did he determine to prove that it wasn't. A few days later he conceived the idea of the chain reaction for which he filed his patent application of 1934. Historians may describe the epiphenomena, but the root

cause of the nuclear age, it would seem, was Szilard's itch to prove Wells a prophet and Rutherford an ass.

Against such anecdotal Hilts describes with feeling the aesthetic and practical sense which Wilson brought to his great creation, the building of Fermilab. From that great cathedral on the Illinois plain miracles may come, one day. On one occasion a friendly Senator, seeking to justify the expense, pressed Wilson to describe how Fermilab contributed to the national defence. Wilson persistently but politely declared it made no such contribution. At the third prompting, he delivered the fine reproof, "It has nothing to do directly with defending our country except to make it worth defending".

Hilts uses the vehicle of Mark Ptashne, part *enfant* and part *terrible*, to recount the tumultuous public history of DNA from Asilomar to Wall Street. Harvard's ill-fated idea of setting up a company in conjunction with its professors is particularly well described. The subject of the final essay, John McCarthy of Stanford, the son of a Boston labourer, has made important contributions to the field of artificial intelligence. Hilts's method of blending science and scientist reaches its acme in this subtly drawn portrait.

Scientific Temperaments is presumably aimed at the general reader but contains much that will be fresh to many scientists. There is, however, lack of conclusion as to how these three vignettes should be interpreted and what, if anything, they have in common. A brief preface suggests that they show "a clear play between science and personality", and that passion, not dispassion, is a trait that many scientists share in common. If there is any room for cavilling at this excellent piece of writing, it is that such interesting assertions deserved fuller treatment. □

*Nicholas Wade is a Member of the Editorial Board of the New York Times. His most recent book, written with William Broad, is *Betrayers of the Truth* which was published earlier this year.*

Chemistry and Biochemistry of Legumes

Edited by S.K. Arora

There is a continued world-wide interest and research concerned with the exploitation of existing legume crops and the development of new crop plants within this family. This book describes the biochemistry and chemistry of these plants and emphasises practical aspects such as improving nutritive value. £19.50 boards 376 pages illustrated

Fungal Technology

The Filamentous Fungi Volume IV

Edited by John E. Smith and David R. Berry

The use of filamentous fungi in industrial processes is now well recognised and is gaining even more importance with the current awareness and rapid expansion of biotechnology. This volume concentrates on the technology of growing fungi particularly in large-scale liquid fermenters as well as in solid-state cultivation. £32.50 boards 416 pages illustrated

Plant Surfaces

B.E. Juniper and C.E. Jeffree

This book describes the biochemistry, structure, function and physiology of the layer that covers the entire cellular tissue of plants including the roots. One of the main functions of this layer is to protect the plant from attack by disease-causing organisms. £5.25 paper 104 pages illustrated

Studies in Biology

A catalogue is available on request from the promotions department.



Edward Arnold

41 Bedford Square
London WC1B 3DQ

Circle No.09 on Reader Service Card.

Introducing the star of the backyard

Simon Mitton

The Sun, Our Star

By Robert W. Noyes.

Harvard University Press: 1983. Pp.263.

\$20, £16.

THE Sun, a run-of-the-mill star, is halfway through its life. Furthermore it is in the middle range of size, temperature and mass for normal stars. Since the next nearest star is 275,000 times as far away, astronomers have observed our Sun with a curiosity transcending that for the other stars, all of which are mere specks of light in conventional telescopes. The Sun is the only star where we can study a photosphere, with its attendant sunspots, flares and active regions, on a continuous basis. In recent times spacecraft orbiting the Earth have made detailed measurements in the optical, ultraviolet and X-ray regions in a way that is quite impossible for the other stars.

Professor Robert Noyes is an astrophysicist at Harvard University who has specialized in solar astronomy. His book, an introductory account of solar physics and astronomy, is first-rate. The story opens in a way that will immediately engage the target audience of undergraduates and amateur scientists; we are taken to spend an imaginary day at one of the world's finest solar telescopes, at the Kitt Peak National Observatory, Arizona. The account continues in a straightforward way by describing the physical regimes encountered within and around the Sun. The author keeps hard science in the background; instead of equations and derivations there is good and clear prose, used effectively throughout. Noyes adds interesting historical touches where appropriate, and these lend perspective to the narrative. Solar studies have several times led to important advances in physics, for example the discovery of very highly-ionized atoms and the basic mechanism of energy creation through nuclear fusion.

Recent discoveries treated include the rotation of the core, the invention of solar seismology, the value of Antarctic observatories where the Sun can be watched for days on end, and the solar neutrino problem. The neutrino flux at the Earth is expressed as 3×10^{13} neutrinos/sec per reader, which somehow graphically summarizes the difficulty of detecting them. Enduring, but false, stories are laid to rest: no more acoustic heating of the corona — magnetic heating is the answer! The book closes with topical chapters on climate, solar energy and solar evolution. There is no bibliography.

The Sun, Our Star is an excellent introductory account of solar astrophysics and a worthy addition to the long-established Harvard list in astronomy. Just possibly

the achievements of space astronomy should have been more closely woven into the main account. The reader waits until the final one-third of the book for much of this (but has been royally entertained meanwhile). Readers outside the United States might also want to quibble with the confusing mixture of units: °C, °K and °F; meters and miles; cgs and SI; all of these jostle for attention. Diagrams and photographs are plentiful, clear and relevant, and they have the benefit of long, self-explanatory captions where appropriate. This is a fine book that will stimulate and interest anyone who wishes to have an insight into the scientific knowledge of our very own star. □

Simon Mitton, editor of the Cambridge Encyclopaedia of Astronomy, is a science publisher and a Fellow of St Edmund's House, University of Cambridge.

Rocketed to success

Lee A. DuBridge

JPL and the American Space Program.

By Clayton R. Koppes.

Yale University Press: 1983. Pp.320.

\$19.95, £16.95.

THE Jet Propulsion Laboratory is today a huge research and development facility with 4,000 employees, located about six miles from the campus of its parent institution, the California Institute of Technology in Pasadena. It is justly famous for its role in space exploration and for being the first laboratory in the United States to initiate, as early as 1936, a fundamental study of the science and technology of rocket propulsion. This effort, which initially aimed at very modest goals (e.g. a sounding rocket), led to a 45-year succession of achievements in rocketry, missiles and unmanned space vehicles. Hence the appearance of the first — but (as the author insists) “unofficial” — history of this great institution is a notable event.

JPL also happens to be one of a dozen or so peculiarly American establishments, known as Federally Funded Research and Development Centers (FFRDCs). Each of these is operated by a university, or a consortium of universities, under a contract with a government agency which supplies the funds and the overall direction of its goals and operations. In each case the contract involves a triad of parties: the operating institution, the laboratory itself and the sponsoring agency. It is understandable that many administrative, financial, managerial and even political problems arise in such an arrangement. It is to these “societal” problems that Dr Koppes chooses to give much attention. With extensive documentation, including 32 pages of bibliographical notes, he weaves together the complex story of JPL's ac-

tivities and problems. He hopes that the story of one FFRDC might be of interest to others, and will also cast light on the whole area of government-university relations.

This is a worthy aim, but I, personally, feel that Dr Koppes' account is marred by his political-ideological viewpoint, embodied in his assertion that the United States is a “National Security State”. By his own definition, this implies that the United States is close to a military dictatorship. But the relations between the government and the American universities during the past 25 years are hardly of the type to be expected under such a regime. The US Space Programme and JPL have been under the auspices of a civilian agency which had no military mission and carried on no secret projects. National “prestige” — to be attained through scientific and technological accomplishment — was certainly a major goal, but this is not the same as “national security”.

This reservation apart, Dr Koppes has given us a fascinating account of JPL's history. His story of the early days of JPL is especially interesting, partly because he includes, here and elsewhere, material which has not been previously published. The first studies of rocket propulsion began under the wise and kindly leadership of Theodore von Karman, the distinguished aerodynamicist and Director of the Guggenheim Aeronautical Laboratories of the California Institute of Technology (GALCIT). His associates were a few graduate students and enthusiastic volunteers, led by a perceptive and sensitive student named Frank Malina. Having heard of earlier rocket experiments, particularly those of Robert H. Goddard of Clark University, this group was anxious to proceed with rocket research. Karman cautioned them, however, that first they must gain a thorough understanding of the basic theory of rocket combustion and propulsion. Then they should make well-instrumented static tests and, only after that, should they try flight testing. By following this procedure over the years, GALCIT, and shortly afterwards the JPL, became, in Koppes' words, “the seed bed of American rocketry”. However, the GALCIT project might well have remained both struggling (for funds and manpower) and small, had it not been for a series of national and world events which radically changed its course.

First came the Munich crisis of 1938, which shocked the US military leadership into action. General H. H. Arnold, Chief of the US Army Air Corps, who had seen the GALCIT experiments, offered a “huge” grant of \$10,000 to finance the development of JATO (Jet Assisted Take-off) rockets for aircraft. More funds soon followed, and JPL itself now “took off” on the first of its series of major development programmes. Important advances in the technology of both liquid- and solid-propellant rockets resulted in JATO's ultimate success.

The 1943 reports of the German rocket activity led the US Army Ordnance Department to enter into a contract with Caltech for the continued operation of JPL, under which JPL would develop artillery-type ballistic missiles. This effort continued into the Cold War period, and from it emerged the Corporal E liquid-propellant missile and the Sergeant solid-propellant weapon, both of which were eventually deployed in Europe. The latter was the forerunner of the modern Pershing missiles. The anxiety of the government to get these weapons into production as quickly as possible caused much confusion and frustration at JPL. Yet in this period the laboratory learned much about system analysis and development, and also took the first, crucial steps towards creating what ultimately became a spectacularly successful system of electronic guidance, control and communication for missiles and spacecraft.

This missile research was terminated with the appearance of Sputnik and the initiation, in 1958, of the US space programme, managed by a new civilian agency, the National Aeronautics and Space Administration (NASA), which took over the JPL contract. Caltech and JPL were glad to close their secret military weapons programme and to engage in a more scientifically rewarding enterprise. Spectacular

managerial problems among the administrators within the Caltech-JPL-NASA triangle, all of which are fully described by Dr Koppes. This was the era, however, in which the appellation "National Security State" seems to have the least relevance, even though the Caltech-JPL-NASA relations were at their most difficult. The problems stemmed, not from security considerations, but from the policy of NASA to monitor most critically every aspect of the performance of JPL, at a time when that establishment was tackling the most challenging problems in its history.

It was ironic that just when Viking and Voyager projects brought JPL to the peak of its prestige, the government began to cut back on planetary exploration, partly because the Space Shuttle programme was soaking up such a large proportion of the NASA budget. JPL suffered serious setbacks when the US portion of the International Solar Polar mission was cancelled, as well as the proposed fly-by to Halley's comet. The only new project remaining for JPL is the Galileo mission to Jupiter, carrying a probe to be released into Jupiter's atmosphere.

Galileo, the future space telescope and other scientific instruments are to be carried on the Space Shuttle, along with a variety of secret military payloads. This im-

Schemes and dreams of fusion

Philip Davenport

Fusion: Science, Politics, and the Invention of a New Energy Source.

By Jean Lisa Bromberg.

MIT Press: 1982. Pp.343. \$30, £24

CONTROLLED thermonuclear research (CTR) has as its primary goal the demonstration that nuclear fusion reactions between light elements can provide us with a new and benign energy source. But despite over 30 years of steadily expanding endeavour by many countries, now freely collaborating, this goal still eludes us. Opening the first International Conference on the Peaceful Uses of Atomic Energy in 1955, when the subject was still classified as secret by the nuclear powers, Homi J. Bhabha of India predicted, "a method will be found for liberating fusion energy in a controlled manner within the next two decades". Much the same might be said today.

This situation, though favourable for the job prospects of plasma physicists, puts the historian in a quandary; any account of a national fusion programme must necessarily be open-ended — should one wait for a successful outcome to round it off neatly? The answer is an actuarial no; several pioneers of fusion have already died — Thomson, Cockcroft and Chick in Britain, Kurchatov, Sinelnikov and Arsimovich in Russia, and Ruark, Christofilos and Tuck in the United States. (Fortunately Dr Bromberg started her work in good time to interview Tuck.) Others candidly admit to having memories that are dulled or, more embarrassing, inventive; the historian has little time to spare.

As the subtitle of her book suggests, Joan Bromberg has concentrated on the political and organizational aspects of the Magnetic Fusion Program in the United States, while providing just the right amount of scientific background and detail to enable the non-specialist reader to understand the principles of the various experimental approaches described. The result is a fascinating account of how the development of one government-sponsored research project has been shaped by scientists, politicians, industrialists and world events.

Experimental work on CTR began quietly in Britain in the late 1940s, with a total of about a dozen scientists working in three small groups at the universities of Oxford, London and Liverpool. Although this work received no publicity, quite a few physicists knew about it; significantly, these included James Tuck, working at Oxford, and Klaus Fuchs at Harwell. Tuck moved to the United States in 1949, became naturalized, and initiated the fusion programme at Los Alamos which later became

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

February 1958, just after Sputnik — Richard Nixon (then Vice President of the United States) examines a model of Explorer 1, the first satellite built by the United States. On the right is William Pickering, Director of JPL at that time. Lee DuBridge is on the left.

achievements followed in quick succession. JPL built the upper stages and instrument package for the first US satellite, Explorer 1. Then came a series of Rangers, Surveyors, Mariners, Vikings and Voyagers which gave the first close-up views of the Moon and all the planets out to Saturn. One Voyager is now on its way to Uranus, still transmitting data to Earth after six years and over 10^9 miles of space travel. The pictures taken and the scientific information gathered by these missions have given the world a wholly new conception of the Solar System.

These successes were, however, laced with many difficulties and frustrations. There were frequent disagreements on

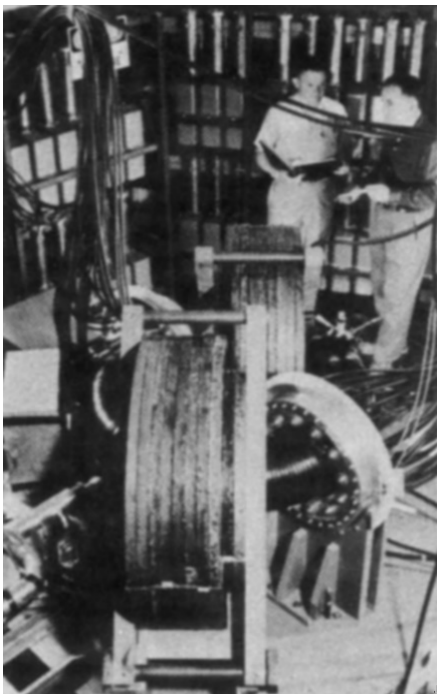
pels Dr Koppes to suggest that NASA is now a part of the national security apparatus, which is like saying that a trucking company is a national security enterprise because it may sometimes carry a military load. Be this as it may, it is clear that both NASA and JPL are undergoing a change in course, and what the future will be is still uncertain. The fact remains, however, that JPL has made scientific and technological history. Many will be glad that this history has now been preserved by Clayton Koppes.

Lee A. DuBridge, a member of the President's Science Advisory Committee from 1969 to 1972, was President of Caltech from 1946 to 1969 where he is now President Emeritus.

Public Information Office, JPL

known as Project Sherwood. Fuchs took a different direction. In March 1950 he was convicted of spying for Russia, and towards the end of that year CTR in the UK was classified as secret, ostensibly because of its military potential. In my view the real reason was an obsession with secrecy for its own sake, to which the United States Government was not immune, for their fusion programme was born classified in spite of concerted opposition from the scientists concerned. Even the very fact that the Atomic Energy Commission was sponsoring CTR was kept secret, as was the location of the experiments, although all three countries must have been aware of each other's interest in CTR. Surprisingly Russia was the first to publicize its fusion research through the famous Kurchatov lecture at Harwell in 1956, much to the discomfiture of the other two.

Subsequently the security barriers were gradually dismantled. First came the UK-US exchanges of information and laboratory visits. Then, after some hard bargaining in an atmosphere of rivalry and



Los Alamos's Perhapsatron S-4, which together with other examples of fusion technology was displayed at the second Geneva Conference in 1958.

suspicion fomented by the Press on both sides of the Atlantic, agreement was reached on the publication in *Nature* in January 1958 of a group of papers which in effect formed a joint progress report on CTR. Bromberg's account of this episode may upset the more sensitive members of the Harwell team even now, and indeed it does contain some technical inaccuracies concerning ZETA, the huge toroidal gas discharge which was Britain's major fusion experiment. But in the context of a history of fusion within the United States, it is in my view a valid and objective assessment of a precarious period in UK-US collaboration.

Complete declassification of CTR was agreed in time for the second Geneva Conference in 1958, where all nations having fusion programmes would be able to reveal the extent of their progress and to exchange ideas. Here the Americans took the opportunity to mount an elaborate working exhibit, the Sherwood Spectacular, which dominated the conference and attracted 100,000 visitors. This was a great boost to US prestige, but it was only achieved at the cost of totally diverting their CTR effort for many months.

The account of the first decade of fusion work in the United States is for me the most interesting part of the book. The four government-sponsored laboratories — Livermore, Los Alamos, Oak Ridge and Princeton — enjoyed a high degree of autonomy. They worked on their individual ideas with a pioneering and sometimes partisan attitude in geographical isolation; a round trip of these laboratories entailed a journey of at least 6,000 miles. This is reflected in the diversity of their early schemes for the magnetic confinement of plasma, some of which bore intriguing names. Los Alamos had its Perhapsatron and Columbus, pulsed discharges in toroidal and cylindrical tubes respectively; Livermore had magnetic mirror traps named Toy Top and Table Top; Princeton's large pretzel-shaped apparatus was called the Stellarator. Oak Ridge's direct current experiment had the more prosaic acronym, DCX. Naming fusion devices gave much scope for whimsical originality; my favourite acronym is Llandudno — large linear and no doubt unstable discharge — coined by Tony Malein at Harwell.

With the advent of international cooperation in CTR, the US experimental programme became increasingly influenced by developments overseas. Over the same period scientific direction and financial control were inexorably withdrawn from the laboratories to the central administration in Washington DC. The later chapters of the book mainly deal with the activities of this central power base under changing styles of leadership and organization. They provide an insight into the vagaries of political priorities and industrial expediency which is instructive and at times cautionary.

This book is, as far as I am aware, the first official history of a national fusion programme to be published. Earlier accounts of fusion research have in the main been written by participants rather than professional historians, and in consequence have been restricted in scope and outlook. Joan Bromberg has set high standards of sensitive objectivity, technical exposition and assiduous referencing for others to follow.

Philip Davenport joined the CTR group at Oxford in 1949 and subsequently worked on early experiments at Harwell and Culham Laboratory. Now retired, he is an honorary consultant to Culham on historical matters.

Style of a physicist

R.H. Dalitz

Tabibito (The Traveler).

By Hideki Yukawa.

World Scientific/Wiley: 1982. Pp.218.

Hbk £22.50, \$38; pbk £11, \$19.

As a person, Hideki Yukawa was little known in Western countries, although he died only two years ago. His name, however, won international renown for his theory that the nuclear forces were mediated by a field due to massive meson, as laid out in his paper of November 1934 published in the *Proceedings of the Physico-Mathematical Society of Japan*, for which he received the Nobel prize in 1949. This autobiography, detailing his life from his birth in 1907 up to 1934 and telling how he came to his idea, has therefore a rather special interest.

The first surprise is to learn that Yukawa was his adoptive surname, taken when he married into the Yukawa family. His father's surname was Ogawa, but that again was an adoptive surname, for his paternal grandfather's surname was Isai. To confuse things further, his father-in-law was born with the surname Sakabe. His father-in-law's adoption by the Yukawa family had a more understandable, although surprising reason, namely that the family Sakabe had been *abolished* by a local Lord, who ordered Sakabe senior to commit *hara-kiri*, so that Sakabe junior was in need of a family.

Despite this unfamiliar background, Yukawa's mind and its course of development conform to patterns quite familiar to us, common among physicists of imagination. The young Yukawa was strongly introverted and his inner life was the more immediate and dominant for him, until well into his late 20s. Much he considers important about his childhood goes against current educational theories. For example, his grandfather taught him the Chinese symbols by rote, introducing him to the great Chinese classics without any understanding of the text in which they were used, a training he counted a great advantage in later life. He also found great value in children's literature of an old-fashioned type which taught him important patterns of behaviour and provided models for his attitudes in life. As he put it, "the chaotic material" which came into his mind in this way "became ordered and organized into life's direction".

Yukawa had no yearning for foreign lands; his life was mostly within him. When Einstein came to Kyoto to lecture, he made no effort to attend (although he was, of course, then only in high school and did not know that he was to become a physicist). However, Yukawa's interest in modern physics was first aroused about this time, by Reiche's book on the old quantum theory and its difficulties, which he found

"more interesting than any novel". Next, Max Planck's volumes on theoretical physics, reaching his hand by chance, conveyed the essentials of the subject to him in a clear and direct way, despite their being in German. His first awareness that physics was on the move at last, came from a lecture given by Professor Nagaoka in a visit to Kyoto during Yukawa's first year as a university student and Schrödinger's papers on quantum mechanics became the young Yukawa's bible in the following year; wisely, he read only from the original papers and came to know them through and through.

Yet, even so, Yukawa felt that he came into theoretical physics, not through conscious choice, but by the exclusion of other possibilities where his personality and intuition gave him no advantage, until theoretical physics was the door most readily open to him. His first postgraduate years were spent as an unpaid research associate in the Tamaki research room at Kyoto University. In 1929, when Heisenberg and Pauli made their famous trip around the world, they came to Kyoto and their lectures were an immense stimulus to the young Yukawa. S. Tomonaga was a fellow student through this period, "an outstanding companion . . . encouraging and challenging . . . the type of person who is aware of limitations and yet comes up with the clever ideas".

In 1932, the discovery of the neutron, the positron and nuclear disintegration directed Yukawa's attention to the nucleus and the forces which held it together. Psychologically the time was ripe for him. Nishina had come to Kyoto about 1930, after a long stay with Niels Bohr at Copenhagen, and in his presence "my solitary mind began to open". Around this time he married Sumi Yukawa, whose photograph had attracted him instantly when a go-between came to his father (Ogawa) proposing marriage. This moved him outside his family circle and the young couple set up house in Osaka, a neighbouring city, as open and lively a city as Kyoto was closed, private and quiet. He was then appointed to a lectureship at Kyoto; Sakata and Kobayashi were students at his first lectures and Taketani came soon after, and they, together with Tomonaga (by then in Tokyo), became his close friends. Next, a university was established at Osaka and he became a lecturer in Professor Yagi's department there; a happy situation, since he felt great admiration for Yagi.

Through this period, Yukawa's work with the nuclear force problem went on almost subconsciously. Already in 1932, he had reached the conviction that there must be some new force field involved but, as he said, "the path thereafter was not straight: I had to take the wrong path first, before I could arrive at my destination". Then, one

night in October 1934, the crucial point occurred to him that the range of nuclear forces must be connected with the mass of the new particle he was seeking. Next morning he calculated its mass to be about 100 MeV. The particle had not been seen experimentally since such high energies were not available in the laboratory. He quickly became confident and predicted that such particles should be found in cosmic rays, in Wilson cloud-chamber experiments. He presented his new theory to the Physico-Mathematical Society of Japan, but his wife kept urging "please

disreputable in those days; physicists were conservative and progress in physics was not to be made in that way. Evidence of a particle was needed first; then we could talk about its properties and purpose! Even in 1954, John Wheeler gave "a sermon" to the American Physical Society with the text from St John: "Believe in those things which are good", urging us to explore thoroughly the established theories of electromagnetism and gravitation, before invoking new physical interactions to account for the elementary particles then being discovered. The resulting work led to

many unexpected and important results, although it achieved little towards our understanding of these particles. Today, a theoretician feels no embarrassment in proposing a new theory involving several dozen unknown particles, nor in proposing some different theory in the following month. This was not Yukawa's style.

Yukawa's autobiography, written in his fiftieth year, is not concerned with his later work, and this helps to focus our attention on the importance of quiet reflection and subconscious

thought on basic problems. We may quote a poem Yukawa once wrote:

In Kamakura,

*here in a deep and quiet valley,
a man walks deep in thought.*

He had in mind a particular professor of philosophy, who impressed him greatly, but we may take it as Yukawa's epitaph. □

R. H. Dalitz is Royal Society Research Professor at the University of Oxford.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

write the English paper and show it to the world"; so, an English-language paper was quickly sent to the Society for publication.

Young physicists today find it difficult to understand the mental attitudes of physicists even 30 years ago. They wonder why it took so long for physicists of Yukawa's time to propose a new particle to solve this problem. The fact is that to hypothesize unobserved particles to account for new phenomena was

Man and machine

Brian Gaines and Mildred Shaw

Micro Man: Living and Growing with Computers.

By Gordon Pask and Susan Curran.
Century, 76 Old Compton St, London:
1982. Pp.222. £8.95.

TO LIVE a life unaffected by computers is becoming increasingly difficult. Soon it may be impossible. Whether the prospect excites or appalls you could well be shaped by reading this book. The medium of print has been used to present the message vividly with pictures and graphics on every page, many in colour. An erudite text has been transformed through brilliant layout and illustration into an everyman's guide to computers. Not only are the techniques here — you can use the book as a guide to chips and bytes, Apples and COBOLs — but also the intuitive feel for the culture.

Gordon Pask should need little introduction. He is an authentic genius, a visionary yet an immensely practical man, whose cybernetic theories have influenced

many in engineering, psychology, education and the arts. Pask built personal computers before the name was coined. Some were intended to interact with people and others to be people. During a quarter-century of intellectual innovation he has used the technology of the day to create models of, and tools for investigating, a range of theories that encompass questions going to the roots of life itself. If anyone can speak for both humanity and computers it is Pask.

Susan Curran brings to the collaboration professional experience of presenting advances in electronic, computer and office technology to businessmen and others concerned with their economic and social impact. Her partnership with Pask has resulted in a lucid book that illuminates the often obfuscated issues of computers and their impact on us, our societies and ourselves.

This is a book of questions rather than answers — it probes and it speculates and it wonders. Many of the questions are profoundly disturbing, not only because we do not know the answers, but also because many of those that we dimly perceive are unacceptable and frightening. Modern

technology has reified science fiction and brought fantasy into our everyday lives. The computer is a tool of the mind allowing us to explore our wildest imaginations; it is also a tool too close to ourselves for comfort, perhaps replacing us in activities that we have been taught are our purpose for existence, such as work.

How can we encapsulate the complexities of computing technology and see it in terms that we already know? One view that has already proved very useful is to regard the computer as providing us with a new communications medium supporting interaction and conversation. Books are remarkable in allowing voices from the past to communicate with us and influence our emotions, values, knowledge and actions. But books are one-way. An author may come alive to us, as may his characters and his concepts, but we cannot answer back or ask questions.

Computers provide a medium analogous to a book that supports two-way communication. The programmer may simulate himself, the personalities of others, and the phenomena of complete worlds of objects and of people that may, or may not, have ever existed. Adventure games on the personal computer in the home allow us to become a character in an electronic "novel". We learn to live in an alien environment, gain friends and resources, and use them to achieve our goals. Expert systems on larger computers allow the recorded mind of an expert, or the composite "mind" of many, to guide others less skilled in complex tasks. In using them you are discussing with a "colleague" problems of medical diagnosis or oil exploration, or the inventive processes of mathematical discovery. You tell them your problem and the information you have. You discuss it with them, query their judgement and ultimately come to a decision based on a collaboration with someone who may be long dead.

Computer technology affects the core of our being because conversation is at the heart of human civilization. We are a species remarkable for our adaptive and learning capabilities, the effect of which is immensely amplified by our capability for conversation. Only one person need learn from experience. Others can be told and media allow the telling to transcend space and time. New media *do* change our world and the computer provides the first one for encoding not just a conversation but the capability to converse itself, not just a picture but the capability to be in the world portrayed. Pask and Curran have made superb use of an existing medium to give an exquisite portrayal of that which is to come. □

Brian Gaines is Professor in the Department of Industrial Engineering at the University of Toronto, and Mildred Shaw Associate Professor in the Department of Computer Science at York University, Ontario, Canada.

Futuristic futures

Margaret A. Boden

The Super-intelligent Machine: An Electronic Odyssey.

By Adrian Berry.

Cape: 1983. Pp.180. £7.95.

The Intimate Machine: Close Encounters with the New Computers.

By Neil Frude.

Century, 76 Old Compton St, London: 1982. Pp.190. Hbk £8.95; pbk £4.95.

THE Duchess assured Alice that "Every thing's got a moral, if only you can find it". The moral of these two books is not hard to find. They both offer a taste (a soupçon) of what has already been achieved by artificial intelligence, but their main message in each case is: "You ain't seen nothing yet". The specific "nothings" the authors have in mind, however, differ.

Berry caps a highly superficial account of some current programs with the forecast that super-intelligent machines will inevitably surpass and replace us. They will outlast all conceivable biological species, which are limited by thermodynamics: long after the Universe is too cold for life-forms to persist or to be formed, intelligent computerized systems could continue to evolve. They could even forestall their own annihilation, by devising a "technological fix" on a literally cosmic scale to prevent the otherwise inevitable collapse of the galaxies into one infinitely massive point.

Berry admits that this is stranger than the strangest science fiction, but his readers' respect for Wells and Asimov may not suffice to settle their doubts. What I found no less difficult to accept is the vulgarity of style and superficiality of argument throughout the book. The question "Will time end, or will it go on for ever?" is given short shrift: the Big Bang theory is cited as showing definitively that time had a beginning, and then our question is answered for us, since "What had a beginning may surely have an end". Again, instead of discussing Abelson's pioneering "Goldwater Machine" (to which he merely refers), Berry tells us that he himself wrote a "Wedgwood Bann Machine", whose quoted conversation provides an easy crack at his *bête noire*, but which casts no light on the way in which computer models of belief systems actually function. An author who is science editor of the *Daily Telegraph* could surely have served his readers better.

Frude's book, which is the more original and interesting, extrapolates from those current programs which embody some sort of "friendliness" and/or "personal" characteristics. These include simulations of political or paranoid belief systems, systems capable of automatic speech-recognition and speech-production (the latter with increasingly realistic vocal timbre and intonation), and computer models

of patterns of social interaction and of distinct personality-types.

The author is a social psychologist, interested in the forms of interaction that may develop between human beings and computers. He argues that — irrespective of whether a programmed system could in principle be *really* friendly — appropriately-programmed computers will increasingly *appear* to be personal beings, and will be so taken by their human users, who will experience interaction with them as a form of intimate relationship. He reminds us that even today's crude beginnings are viewed personally by many people, and that users often adopt an inappropriately anthropomorphic stance towards computer systems. His main point is that this is due not merely to a lack of computer literacy, but to a universal human tendency toward animism, and that this response is more likely to wax than to wane. With increasingly human-like performance on the part of programmed systems, our natural animistic tendencies will often swamp our scientific understanding (which for many users may be minimal in any case).

It is a rare adult who has never smiled back at a smiling teddy-bear, or winced when it was punched on the nose by an unfeeling (*sic*) person. But these sorts of dolls are mere child's play. The novels of Tom Sharpe, if not our own experience, have alerted us to the existence of less innocent sorts of doll: rubber inflatables sold by sex-shops for unprintable purposes. Frude sees it as a commercial inevitability that, among the many friendly and convivial programs to be developed in the future, some will be embodied in "soft machines", sexually seductive or merely cuddly depending on the market concerned. Young monkeys spend more time clinging to a non-feeding monkey-doll that is covered with soft towelling than to the wire-frame "mother" which provides them with milk. Our comparable preference for softness and warmth, combined with our animism, could lead us to have significant emotional engagements with the dolls of the future — whether suited to sex-shops or not.

Past achievements in puppetry, impressive as they are, will be as nothing to the programmed versions. Already in 1970, a soft robot was exhibited in Japan that had 30 distinct "facial muscles" to generate expressions. A wide range of emotions were attributed to this creature by its human audience. Slight changes of speed or muscle-combinations often led to significant

Fresh yeast

Cold Spring Harbor Laboratory have published the companion volume to *The Molecular Biology of the Yeast *Saccharomyces*: Life Cycle and Inheritance* (reviewed in *Nature* 299, 92; 1982). The second volume, Monograph 11B, is subtitled *Metabolism and Gene Expression* and costs \$75 in the United States, \$90 elsewhere.

changes in the interpretations given to the robot's behaviour; for instance, a pleasant smile executed at half-speed appeared as ugly and menacing. Obviously, it would have been a relatively simple matter to provide the robot with artificial tear-glands, the tear-ducts being opened only when the appropriate facial expression was being displayed. Frude points out that future tearful statues of the Virgin could appear even more miraculous than any past attempts at religious trumpery.

"Trumpery" here is my term, not Frude's. In general, he is curiously coy about the moral evaluation of these predicted developments, or at least sees this as more ambiguous than many others would. He forecasts the end of social isolation, when people can buy "a friend off the shelf", and the alleviation of domestic suffering given the widespread use of programmed marriage-counsellors and anti-suicidal Samaritans. Religious and political persuasion and counselling, too, will be available, untouched by human hand. He does not seem to be speaking with tongue in cheek: there is hardly a suggestion that the "social" communication and "religious" counselling he mentions are different in principle from those we enjoy today. There will of course be some deficiencies in practice, due to insufficient advances in the technology; but these, he says, will be largely compensated for by the human users — much as we compensate for the abnormal speech of a foreigner, or of a laryngectomy patient. And he repeatedly rebuts the general question of whether intimate computers *should* be written, always with the claim that irresistible commercial pressures will make them widely available, whether we like it or not.

The general reader may be misled into believing that an impressive artificial intelligence is only just around the corner. Some forward-looking businessman, dazzled by Frude's forecasts of vast profits, may even rush to beg a loan from his nearest merchant-banker — though, bankers being cautious characters, his mission should be fruitless. Neither author sufficiently emphasizes the difficulties involved, the enormous size of the gap between current programs and the human mind. Certainly, Frude does alert his readers to these difficulties every so often. After describing the smiling robot, for example, he points out that "Getting a machine to laugh is easy. Getting it to laugh at a joke is very, very, difficult". Berry, by contrast, writes as though the difficulties were negligible, as though our evolutionary replacement by the first non-biological species — a species capable of influencing the nature of the end of time itself — is a clear certainty. But then, seeing a joke may be difficult for human beings, too. □

Margaret A. Boden is Professor of Philosophy and Psychology at the University of Sussex. Her books include *Artificial Intelligence and Natural Man* (Basic Books, 1977).

Thoughts of sorts

Stuart Sutherland

Science and Moral Priority: Merging Mind, Brain, and Human Values.

By Roger Sperry.

Basil Blackwell/Columbia University Press: 1983. Pp. 135. £12.50, \$16.95.

MANY years ago I worked briefly in Professor Roger Sperry's laboratory at Caltech. I was running experiments on goldfish and on cats, and from time to time he would pop his head round the door, enquire "Have you stopped feeding the goldfish to the cats yet?" and disappear with a chuckle. The chuckle quotient of *Science and Moral Priority* is low, and it is apparent that having won a Nobel Prize for his elegant research on the development of the nervous system and on the split-brain, Sperry has turned his mind to higher things, namely, the nature of consciousness and the selection of moral values. From Charles Sherrington to John Eccles it has been characteristic of neuroscientists to devote their declining years to metaphysical problems. It is also characteristic that they have refused to burden themselves with a knowledge of any previous work on the subject, thus leaving themselves all the freer to speculate. Sperry's new book upholds this tradition.

On the mind-body problem, he adopts a monist position, in which consciousness is an emergent property of the brain. For him, consciousness is made up of complex patterns of activity in the nervous system which apparently can be understood only in mental not physical terms. He believes these patterns of activity are identical to perception, thinking, taking decisions and so on. Since they determine other events in the nervous system and control the muscles, he ascribes to the mind a causal force. He appears to be torn between his scientific view of the brain and his desire to preserve the importance of consciousness, since on page 68 he writes that consciousness is not reducible to neural events whereas on page 96 he states that it is.

Sperry's main and frequently repeated argument for his form of monism, which he modestly describes as a "conceptual breakthrough", is that it constitutes a "paradigm change" that is widely accepted by the informed majority. Leaving aside the danger of deciding truth on a ballot, it is not true that the majority of cognitive and brain scientists, whether informed or not, adopt this view of the nature of consciousness. Most of them are wise enough to hold no view and the fact that the phenomena of consciousness have received increased attention over the past 30 years has nothing to do with anyone's opinions of their status. Indeed, Sperry does not attempt to specify how this "paradigm change" has actually influenced scientific research on cognition and the brain, nor could he do so

since, provided consciousness is in some way linked to brain activity, its status is irrelevant. He is right in thinking that any higher-level system can be explained only in terms of concepts appropriate to that level: the behaviour of molecules cannot be explained solely in terms of the concepts that apply to individual atoms nor can the computations carried out by a program be explained solely in terms of transistors. But this merely reflects a limitation of the human mind, which cannot manipulate a large number of separate entities at once and which must therefore construct higher-level concepts in order to apprehend high-level systems that arise from the interactions of more basic entities. Moreover, Sperry's conclusion that the workings of the brain can only be explained in terms of conscious processes is incorrect. It is at present wholly unclear whether the everyday concepts used to describe consciousness will turn out to be the most appropriate ones to apply to higher levels of organization in the nervous system.

Despite his defiance of William James's sensible injunction that "Dualism is a fundamental datum. Let no man join together what God has put asunder", Sperry is honest enough to admit that his thesis has no empirical support, though he resists the implication that it has no consequences empirical or otherwise. Unfortunately, the only consequences he specifies turn out to be vague promises. According to him, his view of consciousness will "logically dissolve . . . the classic fact-value and naturalistic fallacies of philosophy" (whatever they are). Sperry believes that his identity theory of consciousness and brain states (which incidentally is hardly new) will not only make philosophers redundant, it will allow brain scientists to take their place, for the influence of his theory has apparently been such that neuroscience is no longer "dehumanizing": hence the way is open to found human values "on a reference framework based on empirical evidence and the scientific method". He gives two main arguments for this curious view.

First, he believes that since science has proved the most successful method for discovering empirical truth, it must also be the best way to select moral ends. Even the premise of this non-sequitur may be doubted, since there are many kinds of truth and it could be argued that literature has implicitly revealed much more important truths about human values than has the scientific study of man. Moreover, Sperry does not disclose how science can reveal moral ends. He adopts, without acknowledgement, a rather broader version of C.H. Waddington's view that it is morally right to promote the ends (whatever they are) of evolution. Sperry talks about the "upward thrust" of evolution, without realizing that the word "upward" implies a value judgement that cannot be derived merely from the scientific study of evolution. He seems to include in evolution

not only the animal kingdom, but the planetary system and the cosmos itself. He provides only one example of a moral value derived from science and that is the preservation of the biosphere. Although it has taken science to discover — and to some extent create — the threat to the biosphere, anyone, however unscientific, who is concerned about the future of life will readily agree that the biosphere should be preserved. Scientific findings may both point to dangers and provide means by which conditions can be either ameliorated or made worse, but they are irrelevant to the question of ultimate ends. It is unclear, for example, how science would resolve such ethical problems as abortion; Sperry himself notes that it is a way of reducing overpopulation, but this oversimplifies the problem and there are many other ways of reducing population growth.

A second reason he adduces for the importance of science is that "a science of values in the context of decision theory becomes conceivable". He hints that scientists may one day find methods of forming and controlling people's values. Any such development, which fortunately is extremely remote, is likely to create more moral problems than it solves. Sperry does not stop to consider who would control the use of such methods or the consequences of their being abused. Nor does he recognize

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

that much of the trouble in the world is created not because people have wrong values, but because they fail to act in accordance with the moral values they have. Indeed, so great is his enthusiasm for science that he ignores all the problems that fundamental discoveries about ourselves might bring. Where there is a plausible causal account of an action (for example, if it can be attributed to schizophrenia or a brain tumour), we no longer hold the agent responsible. What would happen to the idea of responsibility if Sperry's science of values were sufficiently developed to explain in causal terms the values held by

everyone?

It will be apparent that this is a disappointing book. It consists of disconnected essays, whose tiresome repetitions serve the place of argument. There are too many arguments by assertion, too many unsubstantiated and unillustrated claims, and too many vague promises. Although Sperry casts himself in the role of a prophet, he is disappointingly short of revelations. Moreover, he has a ponderous style that conceals what he is trying to say: he writes, for example,

This derives in part from the conviction that the functional organization of our cerebral machinery is intrinsically such that the scientific method offers the most reliable means by which

a brain can arrive at an operationally valid stance in the realm of values, as elsewhere.

I think this means "I believe that because of its organization the brain can best discover true values by using the scientific method". To borrow some of his favourite words, his book presents a holistic, mentalist, nonreductive, emergent, dynamic, configurational perspective, but, for myself, I cannot help looking back with nostalgia to the days when Sperry's concerns were with more practical matters, like feline ichthyophagy. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Trouble at t'lab

Walter Gratzer

Betrayers of the Truth: Fraud and Deceit in the Halls of Science.

By William Broad and Nicholas Wade.
Simon & Schuster: 1983. Pp.256. \$14.95.

STORIES of heroic deeds in the laboratory act on the reading public like chloroform. Science becomes news when a microbiology technician catches bubonic plague or a professor is set on by a rampaging mob of anti-vivisectionists. Scientific fraud is a subject that has featured widely in science reporting of late, helped by the ill-suppressed embarrassment and indignation of the scientific establishment at the intrusion of the outside world into a private grief. The description in the opening chapter of *Betrayers of the Truth* of the inept attempts of a group of scientific mandarins, brought before a congressional committee on the subject, to get the toothpaste back in the tube is sheer bliss.

William Broad and Nicholas Wade have put together a most captivating book. We are most of us *voyeurs*, and nothing engenders so much prurient fascination as a gross and indecent violation of the rules by which we are expected to regulate our working lives. Throw in the accompanying discomfiture of stout and dignified parties in positions of authority, and you have all the components of what makes news on a wider stage.

Broad and Wade's purpose, however, is more than just to entertain, for they believe that a close study of the pathology of science will reveal aspects of its normal function that might otherwise remain hidden. Their thesis is that the scientific process is no more informed by reason and objectivity than are other human occupations, that it owes more than is generally conceded to cultural prejudice and rhetorical convention, and that even the best scientists employ in the pursuit of truth a mixture of motives and methods, many of which may not bear close scrutiny. This is not of course the first time that propo-

sitions of this nature have been expounded, but Broad and Wade go on to argue that the conventional wisdom about how science works is shown up for a shallow and meretricious conflation by the response of the system to fraud. They are concerned particularly to refute the often heard assertion that falsehood is rapidly eliminated through its failure to meet the criterion for all publishable work, namely that it must be capable of reproduction in other laboratories. In most of the case histories given in the book, the fraud was indeed not exposed by this mechanism. Most commonly it was uncovered accidentally, but more often, it seems to me, than Broad and Wade allow, a gamey smell issuing from published descriptions of the results was picked up by sensitive noses early enough to save expenditure of useless effort.

Leaving aside a number of now celebrated (though still perplexing) historical cases, in which discoverers of great truths (Ptolemy, Newton, Mendel and more recently it seems R.A. Millikan) apparently felt impelled to give their experimental observations a furtive helping hand, the frauds in areas of major importance described by Broad and Wade — xenografting, for example, and the mechanism of malignant transformation by a viral gene — tended to be quickly detected. This has not been so in that vast undergrowth of science, in which the sole criterion of accomplishment is publication, and the products are phantom papers in the sense that they are probably read by no one at all and will never be cited. One of the most remarkable cases in the book concerns a young postdoc, who perpetrated a series of exceptionally cool frauds while working at Harvard Medical School, and during his two years there produced one hundred publications. Can anyone doubt that the outcome of such proceedings must at best be trivial, and that it scarcely matters whether conscious deception plays any part in them? One might argue that the same holds for much of the output in the "soft" sciences — prudence deters me from being too specific — in which all results tend to reduce to the basic form "some do, some don't".

Betrayers of the Truth concerns itself not only with fraud, but also with other forms of professional misconduct. Indeed the authors do not make much of a distinction between the scholarly sin against God and those against man such as plagiarism, nor do they exclude mere featherbrained credulity, such as emerged so startlingly during the brief reign of Uri Geller and the infant spoon-benders. There are many ways of defrauding one's fellow scientists, the most common being probably the several forms of skulduggery open to referees of papers and grant applications. There are practically no penalties for such abuses;

Thou shalt not steal; an empty feat,
When it's so lucrative to cheat.

Outright fraud is generally fatal if discovered, though even here there are exceptions. In the only case of which I have first-hand knowledge, the perpetrator became a professor and is still hard at it ten years after his first transgressions came to light. So much for the consequences of breaking the tribal taboos. Perhaps just one ritual sacrifice every few years is sufficient to expunge all our accumulated sins.

The question that Broad and Wade do not attempt to answer is how much outright fraud actually retards scientific progress. The answer, I would hazard, is not very much, even if, as they surmise, there is a great deal more of it about than is apparent — in accordance perhaps with the natural law that seven-eighths of nearly everything is invisible. Fraud no doubt reduces the signal-to-noise ratio a little, but its effect is surely trivial compared to that of honest error. However Broad and Wade argue that the incidence of fraud in modern science is a symptom of corruption at the core. It stems, they suggest, from the trend towards large and impersonal research

many quarters on numbers of publications, which to the ignorant remain a more tangible measure of virtue than quality, and you are left with a rich compost in which unhealthy seeds will surely grow. The actions of many of the malefactors in cases of fraud are usually not those of rational men, and in their revelations from the penitent's stool, they speak most often of intolerable stresses applied usually from above.

I would also hold, however, that in a system that ultimately depends on good faith, no one is proof against an ingeniously contrived fraud and a number of blameless and respected scientists have undoubtedly suffered. Nor indeed would one suggest that circumstances can ever altogether exculpate the miscreant. Why, when all is said, did they do it? I think it was Sutton, the bank-robber, who was asked why he was so given to holding up banks; "because" he replied "that's where the money is".

A final reflection: a disproportionate number of the grosser cases of deception described in the book occurred in clinical departments of medical schools, where the exhortations to get ahead by way of publications appear often to be more insistent, and the critical standards consequently lower. There is also, by Broad and Wade's account, a marked reluctance to root out fraud when it occurs and impose penalties to match the offence. And what has happened, you may ask, to the egregious young men, who led their elders by the nose at Harvard, Yale and a clutch of lesser medical schools? Why, they have been sentenced to forfeiture of federal grants and condemned to drink the water of poverty and eat the bread of affliction in medical practice for a while. Dr Johnson said that a man is seldom so harmlessly occupied as when he is making money. Were I a patient, I think I would sooner these enterprising citizens were back in the

Ethics in embryo

R.G. Edwards

Test-Tube Babies: A Guide to Moral Questions, Present Techniques and Future Possibilities.

Edited by William A.W. Walters and Peter Singer.

Oxford University Press: 1982. Pp.165.

Hbk £12.50, \$16.95; pbk £5.95, \$8.95.

HALF the world seems to be opening test-tube baby clinics, with long waiting lists of patients. This procedure offers a good chance of pregnancy for infertile couples — of whom there are so many — while simultaneously raising deep ethical issues on topics such as surrogate mothers or research on human embryos *in vitro*, involving the most fundamental aspects of human conception. Such issues now need a close and impersonal examination, to provide insight on the "value" to be accorded to the human individual at various stages of life. Clear definitions and attitudes are essential: even the use of the term "embryo transfer" needs to be qualified, as it raises the spectre of transferring an embryo from one woman to another.

It is disappointing then that large parts of this book, the result of meetings held in Australia, are too repetitive of earlier discussions. The opening chapter — the inevitable description of the "state of the art" scientifically, which seems to be a must in every book on ethics — is an exact copy of widely-reported events of five and ten years ago in Oldham. Accounts in other chapters of the despair and hope of patients reflect exactly the widely-discussed dilemmas of those who first joined with Lesley Brown and others in a desperate search for their own child. The contributions on ectogenesis, cloning and hybrids similarly offer virtually nothing new that has not already frightened timid ethicists; and the issue of surrogate mothers has been with us since 1970. Today, clarification is needed on values — how they are attained, the sources of ethical and moral decision, the role of the individual scientist or doctor in society, and of the constraints to be imposed on him by professional organizations and legislation.

Some of the chapters dealing with such topics are highly stimulating. I enjoyed Helga Kuhse, on the nature and universality of ethical judgements, their relationship to religious belief and the application of these principles to fertilization *in vitro*. Brian Johnstone draws attention to the moral status of the embryo; the conflict between the life-fulfilling purpose of alleviating infertility and the life-ending action of studying "spare" embryos pushes him close to an absolutist position. In contrast, Kuhse and Singer accord very few rights to the embryo, virtually none. They reject the principle that being a member of the species *Homo sapiens* confers a right to

Harvard Delays in Reporting Fraud

Six months passed before officials made public a confessed case of fraud, and then only after NIH questioned suspicious data

In May 1981, several young researchers at the Harvard Medical School watched in astonishment as one of their colleagues flagrantly concocted data for an experiment. Here was proof, the stunned eyewitnesses later told their superiors, of something they had suspected for months — that their colleague was

May, it was not until NIH questioned results that Harvard discovered the dearth of raw data for the NIH study. At least one faculty member has criticized Harvard for delays and lack of thoroughness. Not the least troubling question in the Darsee episode is why he was allowed to remain in the lab for 6 months,

the hardest working individual that I've ever come in contact with."

In July 1979, Darsee joined Braunschweig, the Hersey Professor of the Theory and Practice of Physic and physician-in-chief of the Peter Bent Brigham and Women's Hospitals. Darsee worked there for the rest of his career, spending

Fraud in Science, 29 January 1982 — a report by William Broad of the case of John Darsee.

groups, in the thrall of robber barons, whose power and rapacity is seldom matched by the intellectual authority needed to run a team of twenty or thirty junior research workers. It is undoubtedly all too common for the leader of such a group to exert relentless pressure on its members to produce or else forfeit his patronage, or, worse, to encourage them to compete with each other for his favours. Add to this the stress that is still laid in

laboratory accumulating inflated academic capital than wondering how they might turn me into an interesting publication. As an excited professor of medicine once said to me in a gratifying moment of self-revelation: "this patient is definitely publishable". □

Walter Gratzer is in the MRC Biophysics Unit, King's College, University of London.

life (a concept they call "speciesism") and support their case by quoting the acceptance of post-coital contraception such as the use of IUDs in modern society.

Carey considers informed consent in relation to the Nuremberg Code, the Declaration of Helsinki and earlier debates on the right of an embryo to give consent to its own beginnings. Discussions on the rights of embryos to consent have always seemed to me to be astonishingly devious, as if embryos can be or are ever asked anything at all, and I was delighted that in his contribution Carey agrees by quoting Fletcher's "Needs are the moral stabilisers, not rights". Carey, however, sadly passes over one of the most difficult questions: how can parents realistically consent to the use of an embryo for research without making massive decisions about embryonic rights? Surely parents can consent only to therapeutic treatments on an embryo; when used for research it is bound to succumb, and parental consent for this implies that the embryo is their material possession, without any right to its own life.

In later chapters of the book, Daniel and Morgan accept the desire of childless couples for children, and stress that the patients are not depersonalized by the procedure. I have long suspected arguments that conception *in vivo* expresses "love", whereas fertilization *in vitro* debases it. The loving demands of infertile couples, one to the other, during fertilization *in vitro* can be far greater than during a chance sexual act leading to conception. I also appreciated Henley's argument, that the small size of the pre-implantation embryo has little relation to any judgement of its "humanness". Rassaby emerges as very liberal and sympathetic on the surrogate mother, a somewhat unusual ethical stance, and Walters braces up to a difficult assignment on cloning, etc., concluding by condemning any potential studies on man-animal hybrids. I fully agree with him.

Some fundamentals of ethical debate are conspicuously absent from this book, many of them concerned with the science and medicine of human conception. Repetition of the absolutist case of the "camel's nose" argument must simply be rejected: undesirable ends are avoided in many areas of human life and can also be avoided with conception *in vitro*. Many aspects of early human embryology also add support to the rejection of absolutism: the genotype of a surprisingly large number of embryos is not finally established at fertilization — as Hassold has recently reminded us with reference to mosaic trisomies — and no one could give any value to an androgenetic embryo destined to result in a molar and perhaps cancerous pregnancy. Worse, however, is a lack in this book of the details of the enormous potential for good in studying early embryos — identifying those with genetic defects, studying differentiation and normal and abnormal growth, teratogenesis, aspects of cancer, even the use of

embryonic stem cells in medicine — the list is almost endless. The enormous research potential for science and medicine, and consequent benefits for ameliorating mental and physical suffering, is for me decisive in debates on the ethical dilemma. The question of children conceived *in vitro* involves one set of ethical issues; the rights or otherwise of embryos *in vitro* raises quite different arguments about the value of research in relation to ethics.

A conflict arises for many of us in contemplating these research studies. Human life must be accepted ethically as being different from animal life — I cannot accept Kuhse and Singer's "speciesism" charge. Equally, we cannot use the small size of a pre-implantation embryo as a moral arbiter, or accept excuses that embryos can be used for research because they would not be there but for *in vitro* fertilization. The fact that some women wear IUDs does not allow us to evade our personal ethical decisions. Nor do concepts such as pain, rationality and consciousness really help in solving our ethical problems on the early embryo, because such parameters take us far into fetal life — weeks, months or even years after fertilization. Such concepts are unhelpful and misleading. The conflict is worst for those who regard a human embryo as different to, say, a mouse embryo, yet accept the need for embryonic research to alleviate disastrous conditions in children or adults. How can notions such as "respect" or "dignity" for an embryo be applied when it will be used for such a need? Decisions about research are made easily by those who believe the embryo has no rights at all, but not by those with reservations.

There is one other significant omission in this book. Where do we find the viewpoints of the Australian scientists and doctors actually involved in the work? They are in the front line — developing the techniques and as a consequence undeniably enmeshed in the ethics. Walters, the co-editor is a gynaecologist, but evidently not involved in the field directly. After their "state of the art" message, the participants in *in vitro* fertilization leave almost all of the ensuing debate to theologians, ethicists and philosophers. How do such practitioners answer Archbishop Little's question in the book: how they gained "the moral right to embark on a procedure which placed them in the predicament of 'destruction or freezing'." Surely, those most clearly involved should have something to contribute, something others cannot? Such inertia inevitably leads to gibes such as "What can be done will be done", and the implication of an uncaring, unfeeling science is reinforced. Research workers must clarify the issues in debate, publicize their ethical attitudes and justify their work to avoid such implications. □

R. G. Edwards, a pioneer in the techniques of *in vitro* fertilization, is Reader in Physiology at the University of Cambridge.

Talk of DNA

F. J. Bøllum

Man Made Life: A Genetic Engineering Primer.

By Jeremy Cherfas

Basil Blackwell/Pantheon: 1982. Pp.270. Hbk £16, \$15.95; pbk £5.50, \$7.95.

Scientists and non-scientists will be attracted by the title of this book; the subtitle should extend the readership to precocious kindergarteners anticipating the arrival of their first cloning kit from the local molecular biology company. Despite the banner headlines Jeremy Cherfas's appeal to this wide audience will probably depend on individual tastes.

The book presents the cosmos of genetic engineering, beginning with the discovery of DNA structure and the tools required for manipulation of DNA molecules. General procedures for cloning and DNA sequencing are covered, bringing the reader right up to the question of what monster he might like to make. As a word of caution, the social eruption resulting from the regulatory discussions of the mid-1970s is introduced, and then it's back to the *coli* shakers.

Benefits to mankind in the production of rare and important hormones and difficult vaccines, and other biotechnology issues, fill the next three chapters. Benefits received by certain individuals in the form of capital gains are also included for conversational interest. Finally, rethinking the scientific problems already opened wide by recombinant DNA experiments and an outlook for the future brings us to the final page. But who will be there at the end?

Although I believe that recombinant DNA technology is a serious and interesting subject, I had great difficulty getting through this book, perhaps because of the unusual writing style. The story might be performed convincingly on the BBC, but it does not read well. Cherfas's penchant for writing letters to famous scientists and then referencing his own correspondence as documentation of some trivial idea is more irritation than supplementation.

Then there is the choice of subjects for discourse. The author tends to select those topics that he thinks are important because he heard of them from famous scientists. An example is the discussion of restriction enzymes in Chapter 2. While scientific drones may be interested in the general subject of restriction and Type I restriction enzymes, those baby cloners are not. They

Price of Red Deer

In last week's *Nature* (p.734) the hardback dollar price was omitted in the review of T.H. Clutton-Brock *et al.*'s *Red Deer: Behaviour and Ecology of Two Sexes*. Full prices are hbk £20, \$30; pbk £9.75, \$12.95.

want the bottom line — Type II restriction enzymes. Somewhat surprisingly other enzymes are sort of ho-hummed, even though a famous scientist reminded Chérfas that those enzymes are important. Clearly the author is not to be deterred by this kind of fact — he prefers conversation.

Speaking of important facts, there is the description of the preparation of a recombinant phage library (p.113):

Once the phage had grown and amplified their precious cargo several hundred times, the agar was scraped off the plates and mixed with a little chloroform to ensure that the cells that had not yet burst to release their phage would do so. The whole lot, about one third of a litre, was spun in a centrifuge to remove bits of agar and cell debris, and popped into the freezer for storage: a small jar, holding the entire human genome as a collection of handy, ready to use pieces.

It was prose like this that caused me to fail Freshman English composition. And, continuing his interpretation at the undergraduate level of understanding, we find (p.51) that DNA electrophoresis on agarose was a major scientific advance because it avoided adjusting solvent concentrations of column fractions obtained during enzyme purification. And to validate the importance of this last breakthrough we have a direct quotation from a letter to Chérfas (p.52), to wit:

'Agarose gel electrophoresis did not originate with our *Biochemistry* paper', Sharp has told me, 'but was certainly popularized by it.'

There is a lot of silly conversation in this book, and terms such as Frederick's ladder and Walter's cookbook are not the silliest.

One might also question some of the value judgements. There are several references to the tedium of protein purification, and on p.145 we learn that because of genetic engineering of DNA ligase genes 'the tedium of protein purification became much more worthwhile. . .'. Aside from the questionable logic, I imagine that those who purify proteins find it a rather stimulating and useful occupation.

I would guess that my reaction is partly cultural, since Americans expect citizens of Great Britain to write and speak in good English, often mistaking a clipped accent for a sign of intelligence. In the present instance the familiar presentation of questionable facts may make this book popular with academic camp followers, but our new generation cloners and budding venture capitalists will probably look elsewhere for the hard facts of recombinant DNA technology. □

F.J. Bollum is a Professor of Biochemistry at the USUHS, Bethesda, Maryland.

Darwinian delicacies

Colin Patterson

Hen's Teeth and Horse's Toes: Further Reflections in Natural History.

By Stephen Jay Gould.

W. W. Norton: 1983. Pp.413. \$15.50.

REVIEWING Stephen Gould's essays, I was tempted to begin, as he so often does, with a snatch of libretto or a quote from a favourite sage. But I resist the temptation; as with top-class cooking, it may be easy to identify ingredients, but what is put on the table has been transformed in an inimitable way.

Hen's Teeth and Horse's Toes is Gould's third collection of essays, following *Ever Since Darwin* (1977) and *The Panda's Thumb* (1980). As before, almost all of the pieces first appeared in *Natural History*, the monthly published by the American Museum of Natural History. Two of the thirty essays are from other sources, a trifle on the evolution of Hershey bars, and one of a trio on the creationist revival. There is one new piece, a reply to critics of Gould's thesis that Teilhard was in on the game at Piltown.

Gould's themes range from the topical (sociobiology, creationism, a terminal Cretaceous asteroid impact), through the historical (Cuvier, Agassiz, Vavilov), to the timeless 'is Darwinism enough?'. Evolutionary theory is his concern in most of these essays, and he comes at it from all

sides. Gould's own answer to the question about Darwinism is 'no'. He advocates a pluralistic and hierarchical theory, with junk DNA and neutralism active at the molecular level, neo-Darwinism at the intraspecific level and as the explanation for finely-tuned adaptation, but *not* for the origin of species or of all novelty. Species may frequently arise by chance events, and randomly generated species form the raw material of species selection, a process analogous to natural selection, but acting at a different hierarchical level. Evolutionary novelties may not be ground out by slow mill of natural selection, but generated instantly by changes in ontogenetic control, like atavism and homeosis. And the box is shaken now and again by random catastrophes such as asteroid impacts or continental collisions.

Whether or not you agree with Gould, you will surely enjoy and admire the vigour of his style, his agile and fertile mind, and his protean scholarship. Who else could spin a couple of thousand entertaining and provocative words from themes as unpromising as the census, or Steno's views on solids within solids? Only one rival comes to mind, T.H. Huxley, who did the same with 'Yeast' and 'A Piece of Chalk'. I don't know how Steve does it, but I hope he keeps it up. Together with his many other fans, I look forward to the fourth collection. □

Colin Patterson is Senior Principal Scientific Officer in the Palaeontology Department, British Museum (Natural History), London.

ADVERTISEMENT

New titles from Pitman

Introduction to Immunogenetics

M B Zaleski, S Dubiski, E G Niles, R K Cunningham

This self-contained approach deals with three traditional fields of immunogenetics — immunoglobulins, blood groups, and cell surface alloantigens. The information is presented from three interrelated points of view — genetic, structural and functional.

Hardback/191 × 234mm/496 pages

May 1983/ISBN 0 273 01925 2

Price £18.75 net

Temperate Legumes: Physiology, Genetics and Nodulation

Edited by D Gareth Jones and D Roy Davies

Invited papers from a meeting organised by the physiology group of the AAB. The volume will interest plant biologists, plant physiologists, geneticists, microbiologists, agriculturalists and plant breeders.

Hardback/244 × 172mm/448 pages

Sept 1983/ISBN 0 273 08601 4

Price £15.95 non net

Functional Ontogeny

Edited by David McFarland

Starting from basic concepts in quantitative ethology, leading experts in the field of animal behaviour discuss topics relevant to ontogeny — in particular life history strategies and functional aspects of learning.

Paperback/244 × 172mm/256 pages

Sept 1982/ISBN 0 273 08545 X

Price £8.50 non net

Brush Border Membranes Ciba Foundation Symposium 95

Brush border membranes are found in the gut, kidney and placenta. This volume focuses on the structure, molecular organisation and functions of these membranes.

Hardback/299 × 152mm/350 pages

April 1983/ISBN 0 273 79659 X

Price £25 net

Fetal Antigens and Cancer Ciba Foundation Symposium 96

This volume considers evidence that fetal tissues can protect against experimentally induced tumours, and that this protection is immunological and specific in its mechanism.

Hardback/229 × 152mm/272 pages

May 1983/ISBN 0 273 79660 3

Price £25 net

All prices subject to change without notice. Available from your bookseller, or in case of difficulty from **Cashpost Service, Book Centre, Southport PR9 9YF England**. Please enclose remittance with your order and quote the ISBN.

Pitman Books Ltd, 128 Long Acre, London WC2E 9AN England.

Pitman Books

Circle No.23 on Reader Service Card.

In search of the insoluble

Eric Ashby

The Youngest Science: Notes of a Medicine-Watcher

By Lewis Thomas.

Viking: 1983. Pp.266. \$14.75.

NEARLY 60 years ago I enrolled as an undergraduate at the Imperial College in London. I had to pick a subject subsidiary to my main subject and chose botany because I had never even read anything about it. The elementary lectures were given by the head of the department, J.B. Farmer. His lectures so fascinated me that I switched subjects and graduated in botany.

The point of this lapse into autobiography is to confess that in 1923 it wasn't the content of botany that attracted me: it was the way Farmer taught it. Samuel Alexander once wrote that "liberality is a spirit of pursuit, not a choice of subject". It was Farmer's "spirit of pursuit" that caught my imagination. He would be talking about the anatomy of desert plants and suddenly illuminate it with hexameters from Lucretius. From pollination mechanisms he led us to reflect on teleology by way of Aristotle and St Thomas and brought us safely back to neo-Darwinism. What appeared in the College calendar merely as "Botany I" was an unforgettable intellectual experience.

Few teachers have this rare gift. Lewis Thomas, a distinguished medical biologist, is one of them. He was asked to write some light-hearted essays for the *New England Journal of Medicine* (the essays, collected in book form as *The Lives of a Cell* and *The Medusa and the Snail*, were reviewed in *Nature* 258, 24 (1975) and 282, 140 (1979)). He may have intended the essays simply to provide entertainment for medicals. Like Farmer's Botany I, they go far beyond their intention. There is a nice passage in *The Youngest Science* where Thomas tells how he was able to examine the inside of his intestine with a colonoscope. To read his essays is to perform a kind of endoscopy on his mind. They show how some biological event stimulates a whimsical curiosity which grows into wonder and then into a dogged determination to understand. I hazard a guess (Thomas will not mind: he is hazarding guesses all the time in this book) that mycoplasmas appeal to that same region of his mind as Mozart quartets do.

In this book Thomas reflects on episodes in his own scientific career, reminiscences of the kind you might persuade him to talk about in a long summer evening after dinner. His father was a general practitioner, working from home without a nurse or a secretary, in a small country town where patients were treated as people and not clinical material. And this leads Thomas to think about the way modern medical

technology has separated the doctor from the patient, and what a pity this is. For the best medicine the doctor can give for many ailments is a dose of confidence; and that is best administered when the doctor touches the patient. Before the invention of the stethoscope the doctor would put his ear against the patient's heart and chest. The stethoscope came into use,

vastly enhancing the acoustics of the thorax, but removing the physician a certain distance from his patient. It was the earliest device of many still to come . . . designed to increase that distance.

This is no atavistic harking back to the pre-scientific days of medical practice; it is a reminder that although medicine is the "youngest science" and has a dazzling record of achievement, nevertheless there is still much more than science in the treatment of sickness. "If I were a medical student" he says,

I would be more worried about this aspect of my future than anything else. I would be apprehensive that my real job, caring for sick people, might soon be taken away, leaving me with the quite different operation of looking after machines.



Lewis Thomas, "still at work and ready to create insoluble problems if he should ever run short of them".

But Thomas, though he has done much caring for patients in his time, has been compelled, by his inordinate curiosity, to spend most of his career doing research. So the after-dinner talk turns to his war service in Guam and Okinawa. He was sent to deal, among other things, with scrub typhus, but in the event there was no scrub typhus in Okinawa. That meant that he and his colleague were (as he put it) "effectively out of business from the outset". What were they to do with the colony of mice they had brought with them, bedded down under the bunk in reams of toilet paper? Business soon revived: there was a nasty

outbreak of encephalitis. Within a week they had the virus reproducing the disease in mice.

And after the war was over? Thomas wanted what most of us wanted: an opportunity to settle down to ten, twenty, years of productive work. He got it. After two short spells, one in Baltimore and one in New Orleans, he was invited to a chair of pediatrics and medicine in the University of Minnesota. It was just what he wanted: the city, the symphony orchestra, a newly constructed hospital, lively colleagues. "It was the greatest fun, and I thought we would stay there for ever."

But, like a rapidly growing plant, Thomas needed to be re-potted regularly. Four years later he was told he was on the search-committee's list for the chairmanship of the pathology department in the New York University College of Medicine. Would he be interested? His reply tells us a lot about his intellectual ebullience: "if they get down to my name, I'll come in a minute". And it tells us how, despite his successful research career, he had never deviated far from the compassionate interest in patients that his father had had; for a little later he became chairman of the medical division in the Bellevue Hospital,

an ancient set of buildings beginning to fall apart, long open wards filled with the sickest and poorest of New York's citizens, inadequately supported by the city but obliged . . . to take all patients who came to its doors.

His loving but frank account of the hospital is a grim portrait of the attitude to health-care in one of the richest cities on Earth. It provided the "best of all possible medical care for the sickest of patients under the worst of conditions". Every item of supply, from penicillin to toilet paper, in short supply or lacking altogether. Everything — beds, wheelchairs, radiators, windows, elevators, needed repairing. Thomas eagerly accepted the challenge and found himself up against the Byzantine fiscal functionaries in City Hall. A fresh type of problem, but he tackled it with the debonair enthusiasm he brought to the study of endotoxin. ("The key to a long, contented life in the laboratory is to have a chronic insoluble problem and keep working at it.")

Now, as chancellor of the Memorial Sloan-Kettering Cancer Center in New York, Thomas is still at work, ready to create insoluble problems if he should ever run short of them. He would regard *The Youngest Science* as a slight piece of work, as indeed it is compared with his papers on endotoxins, mycoplasma infection and immunology. But the book explains, more vividly than any of the scientific papers could explain, just how those papers came to be written, and what it is like to be a medical biologist. □

Lord Ashby is a Fellow of Clare College, Cambridge, and author of *The Search for an Environmental Ethic*, in Vol. 1 of the *Tanner Lectures* published by the Utah and Cambridge University Presses in 1980.

Powerful mythology

David W. Pearce

The Price of Nuclear Power.

By Colin Sweet.

Heinemann Educational: 1983. Pp. 107.
Pbk £3.95.

IN THE space of only five years the focus of public argument about nuclear power in the United Kingdom has shifted from the all-embracing issues of social cost — accidents, routine radiation, civil liberties, proliferation and the “distance” of decision-making from the voting public — to economic cost comparisons. The 1977 Windscale Inquiry into the construction of a further fuel reprocessing plant established that all the social arguments were at least on the agenda. The 1983 Sizewell Inquiry into the construction of a pressurized water reactor is hearing a great deal about the unit costs of nuclear as opposed to coal-fired power stations. The shift in emphasis is partly because Windscale was concerned with one aspect of the fuel cycle, reprocessing, which has international dimensions. Sizewell is concerned with one new power station, although to make any sense at all it must be the test case for a series-ordering of PWRs on a limited scale over the next decade or so. But much of the change reflects more of a holding operation on the part of the nuclear opposition: the social-cost arguments may not have been entirely won but they have become legitimate. It remains to knock the purely economic costings on the head and, intellectually, anyway, nuclear power would be finished.

Colin Sweet's book is written in this context and it has quite clearly been designed to coincide with the Sizewell hearings, so much so that readers unfamiliar with nuclear arguments in the UK will find it oblique in its institutional references. It is also meant to be part of the culmination of a sustained effort in the past five years to get the economics sorted out. To this end, reports by the House of Commons Select Committee on Energy, the Monopolies and Mergers Commission, reflective views of the independent economists and the pressure applied by the committed lobbies have led to a situation in which a full reappraisal of nuclear power economics should be possible. Alas, while openly written from the anti-nuclear standpoint, Sweet's book does not offer that reappraisal and manages to extend and confuse some of the mythology into the bargain.

Sweet's thesis is, first, that the role which nuclear power *can* play in the energy future is limited to base load electricity which is unlikely to include space heating. This is because consumers do not reveal a preference for substituting electricity for gas or even oil, and because, anyway, the level of *overall* energy demand has been greatly exaggerated by the forecasters. Second, nuclear power is not cheaper than

coal-fired electricity and can only be made to appear so by fudging the books. Notably, what is required is a fully fledged current cost-accounting framework in which investment appraisals, plant by plant, are presented using established appraisal techniques. Third, the chances of demonstrating this last set of propositions are small because of the secrecy surrounding nuclear power, a secrecy resorted to so that committed views on the part of the “pro-nuclear lobby” are not to be revealed as self-delusion.

Sweet is on happiest ground when talking about nuclear versus coal costings (Chapter 5). Even here, however, one would wish for full notes on the sources of data. On the general energy policy context, and in invoking the discipline of economics, he is distinctly uncomfortable and discomforting as a researcher. The problems range from simple mistakes to the setting up of innumerable straw men. Thus Sweet declares that the absolute contribution to electricity output by nuclear power declined from 1971 to 1981. It rose from 19.5 TWh to 23.4 TWh, as his own data show. He declares that total primary energy demand fell by 20 million tonnes oil equivalent over the same period “and continues to fall”, an answer one can only get by careful selection of the base year given that energy consumption rose to 1973, fell to 1975, rose to 1979 and fell in the two years following. These would be quibbles in a short review but for Sweet's conclusion that the Central Electricity Generating Board's case “seems to be somewhat weak” given the context of falling overall demand.

The straw man phenomenon occurs frequently. Because Sir Martin Ryle calculated that nuclear generating capacity would have to rise from 56 gigawatts to a startling 230 gigawatts if *all* fossil fuels were displaced, this is somehow meant to persuade us of the Gargantuan silliness of nuclear claims. But not only would no nuclear enthusiast claim as much, but the quotation of Ryle's computation flatly contradicts Sweet's own very sensible point that electricity of any kind has a limited role in energy substitution in the UK. Much the same disappointment extends to the old chestnut of the Atomic Energy Authority's “programme” in evidence to the Royal Commission on Environmental Pollution in 1976. The AEA is not responsible for forecasts in the UK, nor does it buy generating equipment. Sweet manages to present all of this highly misleading material without once printing the actual official projections of energy demand in the UK. The other errors are manifold, from confusion between rates of return and rates of discount to defining an energy ratio (energy per unit of national output) in terms that can only mean Sweet is thinking of an energy *elasticity* (percentage changes in energy divided by percentage changes in output) whilst consulting data about the former.

The lesson in all this is that we must surely accept the fragility of the relationship between pro- and anti-nuclear power lobbies and, while replete with conspiracy theory, Colin Sweet's book is mercifully clear of the vocabulary-abuse common to both sides. But if Sweet's book is intended as a further nail in what he sees as the coffin of nuclear power he has hammered neither straight nor true. The debate deserves better, especially as those qualified to assess the economics of nuclear power in the sphere of disinterested research are few and far between. □

David Pearce, currently Professor of Political Economy at the University of Aberdeen, is due to become Professor of Political Economy at University College London in October 1983. He is senior author of *Decision Making for Energy Futures* (Macmillan, 1979).

Learning from Seveso

Frederick Warner

The Chemical Scythe: Lessons of 2,4,5-T and Dioxin.

By Alastair Hay.

Plenum: 1982. Pp. 264. \$??, £??

The Chemical Scythe, the first in a series of books to be published by the International Disaster Institute, deals principally with the “lessons of 2,4,5-T and dioxin”. The author, Dr Alastair Hay, has written widely on the Seveso accident. He has now cast his net more widely to examine different incidents involving 2,4,5-T plus dioxin and others in which chlorinated hydrocarbons, such as hexachlorophene, were the toxic agent. This leads him to discussions of the effect of 2,4,5-T used as a defoliant in Vietnam, Laos and Cambodia, and to the possible adverse effects from past dumping of chemical residues in places such as Love Canal near Niagara.

All of these issues raise emotions which vary according to the closeness of the observer. To the British, who often complain of isolationist trends in the United States, there is little appreciation of the intense feelings generated by the war in Vietnam and the claims of ex-servicemen that they have suffered long-term damage from chemical insults. The nature and severity of these insults is witnessed by papers from Vietnam which, however, have not stood up to detailed critical scrutiny by such bodies as the National Academy of Sciences. Dr Hay has set out the various arguments from the literature available, and has added the accounts of journalists who have made their own contemporaneous or later investigations. The reader is left to make his own judgement since Hay gives the opposing views without putting any particular emphasis. He is critical however of the failure of chemical manufacturers to release information in order to help the airing of problems in

public, and concludes that there is still work to be done before the adverse health effects of specific chemicals can be assessed.

Dr Hay could have gone further. Later works from the International Disaster Institute may indeed do this and take up the general nature of disaster, our reactions and the remedies proposed. Unfortunately it appears that analysis of natural disasters will be excluded although they kill tens and hundreds of thousands in floods, typhoons and earthquakes; man-made disasters are never on this scale unless they arise from the failure of dams or embankments installed to control all but the most exceptional floods.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Symbol of Seveso — August 10th, 1976, and workmen begin to cleanse the ICMESA plant.

The special disaster for Dr Hay was that at Seveso, in northern Italy, on July 10 1976, where the ICMESA factory of Hoffman-La Roche made trichlorophenol, used (but not in this case) to make 2,4,5-T, by hydrolysis of 1,2,4,5-tetrachlorobenzene with caustic soda. The reaction is carried out at atmospheric pressure in ethylene glycol solution which is then distilled under vacuum at a temperature not exceeding 185°C. The accident happened following an exothermic reaction, probably initiated from a hot-spot after the agitator and the steam were shut off, so that the temperature rose to 240°C and the reactor pressure to 3.55 bars, rupturing a disk in a vent-line discharging into the atmosphere. The temperature rise caused the formation of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, one of the 75 possible isomers of dioxin, having a toxicity to guinea-pigs and mice represented by LD₅₀ values of 2 and 284 µg/kg. Hay gives an extended account of the chemistry of dioxins, their formation and destruction by incineration, photolysis, electrochemical oxidation and biodegradation, and estimates the amount of dioxin formed at Seveso to have been 250–500 g. This is considerably lower than the 2kg estimated for the Lombardy Regional Council.

The toxicology of dioxin is discussed mainly in relation to animal experiments. The effect on man was shown at Seveso in workers and their families by the rapid development of chloracne, and Hay devotes an entire chapter to this extremely disfiguring disease. The alarming feature of the Seveso discharge was its dispersion over a wide area from the two-phase jet issuing at sonic speed from the vent. Near the source, heavier droplets were caught in vegetation and some animals died after

eating it. Further downwind the smaller droplets contaminated houses and fields, the dioxin content causing 187 cases of chloracne in children but none in adults. These cases had all cleared up by 1980 (the gross lesions on the faces of two children given world-wide publicity were due to the sodium trichlorophenate from the reactor).

Dr Hay records in detail the action taken after the accident by Hoffman-La Roche and Italian government, both local and national. The warnings and clear advice from the company were in contrast with the vacillation of the public authorities; the lack of emergency procedures and political confusion made the people of Seveso victims of a social disaster far greater than any physical damage they suffered from the chemical itself. The publicity given to possible teratogenic effects triggered demands for abortion to be freely available, opposed by the local clergy. Even worse, the agricultural and manufactured products of the area became unsaleable because no firm exculpation came from the authorities. Hay's account shows Dr Reggiani, the medical director of Hoffman-La Roche, acting with the highest integrity throughout the affair. His pursuit in the subsequent years of comparative health statistics for spontaneous abortions, birth defects and many other indicators has shown no detectable differences between the people in the contaminated area and those outside.

At the root of this book is a question — how do we remain alert to the detrimental effects of chemicals shown in some animal experiments without denying to the community products whose utility outweighs the risks? Hay shows how many uncertainties exist; for example, dioxin shows no mutagenicity in the Ames test. Yet the Seveso incident has given rise to the Seveso directive of the EEC, and to the campaign in the United Kingdom to ban 2,4,5-T even though there is already a stringent limit on dioxin content.

It is time for more agreement and less debate on how to compare benefits with costs, particularly on the reduction of risk in the use of medicines. In the Royal Society study group on risk, it was argued that action to reduce the risk of one death would justify spending between £100,000 and £1,000,000, any more being better spent on other medical benefits. In comparison with the estimated cost of smallpox eradication by the World Health Organization it is arguably too much. That campaign has saved 2,000,000 lives a year at a cost of £8 per life. Even more lives could be saved by providing safe drinking water, but the victims are a long way away from Europe and North America. Unlike Seveso and Love Canal, it is hard to make the publicity sensational. □

Sir Frederick Warner is a Consulting Chemical Engineer and Visiting Professor at University College London.

States and soldiers

Richard Langhorne

The Pursuit of Power: Technology, Armed Force, and Society since AD 1000.

By William H. McNeill.

Basil Blackwell: 1983. Pp.406. £15.

FEW AGES have been as preoccupied as our own with the preparations for and likely consequences of going to war. Thus *The Pursuit of Power*, a study of technology, armaments and society as they have developed over a millennium or more, must arouse considerable interest.

The author is William McNeill, an historian who is no stranger to the temptation to seek out and describe patterns of development over very long periods. He has notably succumbed once already in *Plagues and Peoples*, his treatment of the incidence of disease published by Blackwell six years ago. Global histories have not been fashionable in recent times, though there are signs — of which this is an example on a comparatively small scale — that we may not be deprived forever of modern successors to Gibbon and Macaulay. Although *The Pursuit of Power* is not an especially substantial tome, the treatment it offers is far from lightweight. The discussion is tightly packed and marked by great breadth of learning. The conclusions that emerge are often complex, but a finely balanced prose style renders them easily accessible to the general reader.

The pattern perceived by McNeill results from the operation of two factors. First, the diffusion from China during the Sung dynasty (AD 960–1279) of a new kind of economic organization based upon market forces rather than the "command" principle which had characterized the economies of the ancient empires. Secondly, McNeill continues, there was a steady development of ever more efficient weapons, used by increasingly sophisticated armed forces. It was this factor that propelled new economic development. The two effects together came to affect profoundly the creation, structure and behaviour of states. Certainly, it is true that since the evolution of the Italian city states in the fourteenth century, the division of human society into nation states has become a global fact. The more difficult question is whether McNeill's attractively organized explanation for this phenomenon is wholly satisfying.

The ability of economies to generate technological innovation has been a particularly important ingredient of power since the last quarter of the nineteenth century. Economic success, however, and effective military organization have been but two of the ingredients, and it is not certain that a military stimulus has been required on a regular basis to produce successful

technological advance. There have been times when the stimulus was the other way round — particularly, for example, in the application of the steam engine to naval warfare.

McNeill sees military strength as being the principal reason for the internal security of rulers, at least until the middle of the twentieth century. However power may rest with a state for other reasons entirely; internal stability, for example, may follow because a regime has enjoyed an accepted legitimacy. Thus, states have held respected positions more because their systems of government were secure enough to develop effective military forces for external use, rather than because they could dominate their own populations militarily. The eighteenth century provides a classic example of this kind of effect, in which all states improved or wished to improve their internal administration. McNeill thinks that this was just because rulers had come to enjoy a complete internal security through their possession of more efficient armies. It is also true, however, that at that time states were facing an unprecedented problem of international relations in that, because of a shift in the distribution of power among them, they no longer faced a single threat but a multiple one in which each nation was equally feared and fearful. Then, as before and as now, internal efficiency and stability was one of the components of power and it was not surprising that anxious rulers should have attended to domestic reform more because of foreign threats than because they felt militarily predominant within their own borders.

A similar problem for McNeill's thesis occurs in late nineteenth-century Europe. He offers three causes for the First World War: first, a threat to the balance of power (which he regards as the least significant); secondly, changes in the structure of industry and state administrations induced by the demands of military developments, leading to a dysfunction between the striking success of industrial organizations in the advanced states and the increasing inability of those states to apply this advance to their relations with each other. Thirdly, a demographic crisis developed which could only be relieved by war or disease, both of which had occurred by 1918. Here again, while accepting the importance of a demographic factor, we need to be reminded that the chief consequence of technological advances in the nineteenth century, particularly in communications, was a shift in the distribution of power in the world; it was this, rather than anything else that made states feel insecure. The defensive reaction of Germany to the rise of Russia in the early twentieth century exemplifies this effect. In such circumstances all states felt under pressure to make the most of their industrial capacity, their armed strength and their internal stability.

For McNeill, it is the consequence of fighting the two world wars rather than their causes which are profoundly signi-

ficant, because those consequences amounted to an apotheosis of the pattern he has identified. Compelled by military necessities, states had to make fundamental changes in their administrations. This time the heady mixture of danger, patriotism and a formidable level of industrial and technological organization pushed them out of the old path of social and political development, carried on the back of intertwined economic and military change, into a new kind of mixed or semi-command economic and social structure. McNeill argues that since 1945 this situation has been maintained by military and technological pressures. The constant fear induced by nuclear weapons, and the enormous expenditure lavished on armaments generally, have led many to agree that this new structure, and other characteristics of contemporary political life, are militarily determined. Military force, however, is clearly not the only determinant of power and may have more significance as a reflection of its distribution. Some would say that the present baroque, even rococo, state of modern military competition is not just a consequence of the global power structure, but that its complexity reflects the fact that the power structure is less volatile than for many centuries past.

McNeill rightly stresses contemporary international tensions, and he has a prescription for their amelioration: world government, which, as he says, is certainly technically feasible. But the will is so lacking that the very prescription has a quackish air. More sensibly, he observes that it would be wise to expect the future to yield as many surprises as the past, some of them unpleasant. To any historian accustomed to working on any part of the last five centuries or so, the most surprising thing that could happen is that upheavals in human society should slow down and gradually stop. McNeill, however, does foresee this as the likely desirable consequence of the establishment of a world government, and allows that greater demographic control and the immobilizing nature of nuclear weapons have produced greater stability since 1945.

Nuclear weapons may be about to be eclipsed by new weapons but foreseeable technological developments are unlikely to disturb the current balance of power. If this is so, it will be likely to demonstrate that it is not any particular weapon that has been immobilizing, so much as the absence of any serious threat to the dominance of the super powers. They have now achieved an apparently impregnable position against which all the threats of the post-War world have proved ineffective. Perhaps McNeill might be most surprised by the prospect that this fact may eventually permit the natural development of the tranquillity which, for him, can only flow from an attractive but impractical dream. □

Richard Langhorne, an historian, is a Fellow of St John's College, University of Cambridge.

Jewish nationalism and the liberal ideal

M.F. Perutz

The Essential Chaim Weizmann: The Man, The Statesman, The Scientist.

Compiled and edited by Barnett Litvinoff.

Weidenfeld & Nicolson: 1983. Pp. 290.

£16.50.

The Essential Chaim Weizmann conjures up a vivid image of the great Jewish leader, mostly by letting him speak in his own words. It quotes extracts from his speeches, letters and memoirs; it adds summaries of his career, of his political battles with successive Zionist Congresses, and of his brilliant negotiations leading first to the Balfour Declaration and later to recognition of Israel as an independent state. It contains an appreciation, in no more than three pages, of Weizmann's scientific work, and it reproduces a variety of people's sketches of Weizmann's personality. Appendices give the texts of several official documents, from the Faisal-Weizmann agreement in 1919 to Israel's Proclamation of Independence of 1948.

Weizmann had a fervent belief that the Jews were a nation entitled to the rights of a nation; at the same time he was inspired by liberal ideals and averse to any form of violence. His high moral standards can be gauged from his courageous address to the Zionist Congress of 1946, at a time when the British Labour Government was blocking Jewish immigration into Palestine and bitterness was at its height:

It is difficult in such circumstances to retain a belief in the victory of peaceful ideals, in the supremacy of moral values. And yet I affirm . . . that we must retain it. Zionism is a modern expression of the liberal ideal. Divorced from that ideal it loses all purpose, all hope. . . . Assassination, ambush, kidnapping, the murder of innocent men, are alien to the spirit of our movement . . . it mocks the ideals for which a Jewish society must stand.

Already in 1929 Weizmann had written to Einstein: "I have always preached most unpopular realities to the Zionists, have always been attacked most bitterly . . .". For example he told the Congress in 1923:

Palestine is not an empty country . . . there are 500,000 Moslems, 100,000 Christians and a 100,000 Jews . . . there has been a striving on the part of the Arab people for a revival . . . being anxious for a revival of the scattered Jewish people, we treat with respect and reverence any attempt at revival amongst other people.

In his autobiography, *Trial and Error* (Hamish Hamilton, 1949), Weizmann wrote:

We must stand by the ancient principle enunciated in our Torah: "One law and one manner shall be for you and for the stranger that sojourneth with you . . .". I am certain that the world will judge the Jewish State by what it does to the Arabs.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Chaim Weizmann is sworn in as the first President of Israel, 15 February 1949. "We must stand by the ancient principle enunciated in our Torah: 'One law and one manner shall be for you and for the stranger that sojourneth with you . . .'. I am certain that the world will judge the Jewish State by what it does to the Arabs".

There is never a word about the Jews being the Chosen People. He regarded Menahem Begin as "a megalomaniac suffering from a Messianic complex".

Though Weizmann was fond of quoting the prophets, he rarely appealed to God. I gained the impression that he was an agnostic at heart and believed in a secular Judaism as a cultural bond that welds the world's Jews into a single nation. His opposition to bigoted orthodoxy first emerges from a letter to Theodor Herzl in 1903:

If there is anything in Judaism that has become intolerable and incomprehensible to the best Jewish youth, it is the pressure to equate its essence with the religious formalism of the orthodox.

In 1949 he wrote:

It is the new, secularized type of rabbi . . . who will make a heavy bid for power by parading his religious convictions . . . they transgress a fundamental principle which has been laid down by our sages: "Thou shalt not make of the Torah a crown to glory in, or a spade to dig with".

Weizmann was born and raised in the Pale of Settlement "that prison house created by Tsarist Russia for the largest part of its Jewish population". He escaped as an 18-year-old — on a timber barge going down the Vistula — to study chemistry in Germany, determined to raise the Jews to a nation equal in rank with other nations. In his eagerness to awaken Jews all over the world to a national consciousness, and to preserve Jewish culture and heritage in a sovereign Jewish state, he was intolerant of those who have become assimilated to the countries they live in. "I don't care what will happen to the Jewish communities in England or in France . . . their highest ideal is assimilation, disintegration, dissolution". By contrast it seems to me that all the world's Jews could not find refuge in Palestine even if they wanted to, and that by failing to assimilate they will keep antisemitism alive for ever. As Boris Pasternak makes Dr Zhivago say:

Why don't the intellectual leaders of the Jewish people . . . say to the Jews: Don't hold on to

your identity, don't all get together in a crowd. Be with all the rest . . . What the Gospels tell us is . . . that there are no nations, only people.

In the Balfour Declaration the extent of the Jewish National Home had been left open. Weizmann hoped that it would include a much larger area than the present state of Israel. The Turkish province of Palestine included both Cis and Transjordan, and at the Versailles Peace Conference in 1919 Weizmann pleaded that all territory west of the Hedjaz railway, which runs from Aqaba via Amman to Damascus, should be included in the National Home. Today this may sound like imperialism, but judged by the nonchalance with which Lloyd George and Clemenceau disposed of the former Turkish Empire by carving it up between them, Weizmann's claim to the whole of Palestine seems less unreasonable now than, say, France's to Syria and the Lebanon. According to Asquith, Lloyd George backed Weizmann, though he "does not care a damn for the Jews, but thinks it an outrage to let the Holy places pass into the possession of agnostic, atheistic France". However Sir Isaiah Berlin believes that Asquith traduced Lloyd George who did not like the English establishment and had a certain fellow feeling for minorities treated *de haut en bas* by grand Englishmen such as Lord Robert Cecil.

The opening, in 1925, of the Hebrew University on Mount Scopus, for which Weizmann had campaigned since 1902, was one of his life's great triumphs. "Our Hebrew University . . . will have a centripetal force, attracting all that is noblest in Jewry throughout the world". Litvinoff's book might have included more about Weizmann's labours for the other two academic institutions that have made Israel a sturdy outpost of science and learning, the Weizmann Institute of Science and the Haifa Technion.

Weizmann was a phenomenon. He arrived in London in 1904 as a promising young chemist and found an appointment as a research fellow at Manchester University the following year. He proved

himself a highly original and inventive scientist. Thirteen years later he had pledged the British Government to the establishment of a Jewish National Home in Palestine. How did he, an immigrant Russian and self-appointed leader of the world's Jews, achieve this triumph? He did it single-handed, aware, perhaps, of Francis Bacon's dictum: "In all negotiations of difficulty, a man may not look to sow and reap at once, but must prepare business and so ripen it by degrees". For much of the time he was at loggerheads with, rather than backed by, a bickering Zionist Congress. It seems that he was a born statesman whose force of personality made other statesmen accept him on his own terms. In his book *Chaim Weizmann* (Weidenfeld & Nicolson, 1958), Isaiah Berlin has described how

his method of argument was . . . neither a demonstration founded on . . . documented evidence, nor emotional rhetoric, nor a sermon . . .; it consisted in painting a . . . vivid, detailed, coherent, concrete picture of a given situation or course of events; and his interlocutors . . . felt that this picture . . . conformed to their own experience of what men and events were like, of what had happened, or might happen, or . . . could not happen. The moral, historical, economic, social and personal factors were blended in Weizmann's remarkable . . . expositions much as they combine in life.

Yet in the critical years after 1938 Weizmann failed to persuade British leaders to open the gates of Palestine either to the Jews threatened by the holocaust or to its survivors. Torn between his passionate devotion to the Jewish cause and his admiration for Britain, the country that embodied his liberal ideals, he became estranged from both his English and his Jewish friends. This tragedy clouded the last years of his life and is movingly described in Berlin's book, but rather glossed over by Litvinoff.

Weizmann's tremendous moral courage and supreme confidence in his own persuasive powers made him plead his case with political leaders as diverse as Lloyd George, Franklin Roosevelt, Benito Mussolini and Ibn Saud. Ben Gurion wrote of him:

I never failed to be astonished by his inner forcefulness, by his determined manner. He could be angry with these men in power . . . but his anger emerged with such natural dignity that it was always deeply moving.

Lord Robert Cecil wrote of his subdued enthusiasm and the extraordinary impressiveness of his attitude "which make one forget his rather repellent and even sordid exterior". General Smuts, another statesman with whom Weizmann made friends, wrote this epitaph: "He was the greatest Jew since Moses". I wish that the ideals for which he stood would still guide Israel's rulers today. □

M.F. Perutz is at the Medical Research Council Laboratory of Molecular Biology, Cambridge, UK.

Liberal development

P.D. Henderson

Economic Development: Theory, Policy, and International Relations.

By Ian Little.

Basic Books: 1982. Pp.452. \$20, £12.

DURING the quarter of a century that ended in 1973, the world economy grew at a rate which was both unprecedented and unforeseen. Moreover, the rise in output and average material prosperity was not only more rapid than in the past, but more widely diffused. In terms of output, it is the growth of developing countries in particular which has distinguished the post-War era. Furthermore, in most of these countries the momentum of growth was well sustained for several years after the turning point of 1973, when the West succumbed to apparently chronic "stagflation".

On the world political scene the developing countries, their numbers swollen by the dissolution of all but one of the past colonial empires, have become established as a distinct and surprisingly cohesive pressure group. Since 1961 they have been organized for negotiating purposes in the so-called "Group of 77", which now includes some 120 members. As such, they have consistently pressed the rich industrial countries, through the medium of what is now called the North-South dialogue, to agree to a package of international economic reforms. The purpose of these is to redress what is seen as a bias of the existing system against the developing countries, which by holding back their progress perpetuates international inequality and injustice. Since 1974 the Group of 77 have argued for a "New International Economic Order". However, as Ian Little points out in his remarkable new book, the main ingredients of this programme go back as far as 1946, to the early post-War discussions on the form of the international economic system.

A related post-War phenomenon has been the emergence, within the world of academic economics, of a new field of study. Its separate status has been a matter not just of geography, but of underlying methods and assumptions. A consensus emerged that a different approach, a distinctive model or framework of thought, should be applied to the developing countries. To construct this framework, and apply it to particular countries and problems, has been seen as the special task of development economics.

Third age

The final volume in the trilogy *Aging: A Challenge to Science and Society*, has been published by Oxford University Press. Subtitled *Behavioural Sciences and Conclusion*, the book costs £40. For review of Vols 1 and 2 see *Nature* 296, 179 (1982).

All three aspects of recent history — the record and its interpretation, the effects on the developing countries of the international trading and monetary regime, and the evolution of economists' perceptions of reality — are brought out in Little's study. In his preface, he describes the book as "a rather comprehensive, critical survey of development economics", but in fact it is much more than that. Rather, it is an extended and masterly personal review of the leading issues in economic development, development economics and international economic relations. Its unity derives not from a particular emphasis on the ideas of economists, but from the author's own view of the world. The book explicitly sets out a "neo-classical" or "liberal" conception of development, which embodies both a set of values and a vision of the economic process. As to the values, development is defined in terms of the traditional objectives of liberal economics, namely material prosperity and a more equal distribution of income and wealth. Thus it is the welfare of individuals, both living and unborn, that is to count. This goes with a concern for personal liberty, as an additional though related value.

As regards the working of economic systems, the neo-classical approach sees markets and prices as providing, at least potentially, appropriate signals for people and organizations. A market economy serves a dual cause. It safeguards freedom of choice, and it makes for a more efficient allocation of resources. Economic liberals are therefore sceptical about the benefits claimed for centralized planning and administrative modes of regulation and control. To use Paul Streeten's not entirely respectful term, they are "price mechanists".

This approach is seen as generally valid, for rich countries and poor alike. One corollary is that as a rule countries will benefit from the adoption of liberal trade and payments policies, by themselves as well as others. What Little calls the "resurgence of neo-classical economics after 1960" in relation to issues of development mainly arose from two related lines of thought. One was a restatement of the case for free trade in a form which is more precise and less open to qualification. The other was a growing realization, based on observation of and research into the development process — research in which Little himself played a leading part — that countries with open trade regimes had gained from this choice, while those with restrictive policies had suffered. In the former category, the outstanding cases are Hong Kong, South Korea, Singapore and Taiwan. These have achieved not only amazingly high growth rates but also a reasonably even distribution of the fruits of growth. In part at least, this can be attributed to the adoption of open trade policies, which have made the process of growth not only faster but also more labour-intensive.

The liberal approach is at variance with

three other bodies of doctrine which are surveyed at length by Little, often in severely critical terms. One is the orthodox development economics of the 1940s and 1950s. A second and more recent growth is the "new radicalism", in both Marxist and non-Marxist forms. Finally, there is the stance taken by the Group of 77 and their Western sympathizers, such as Willi Brandt and Edward Heath. This forms part of the subject-matter of the final section of the book, which deals with international economic reform and North-South relations.

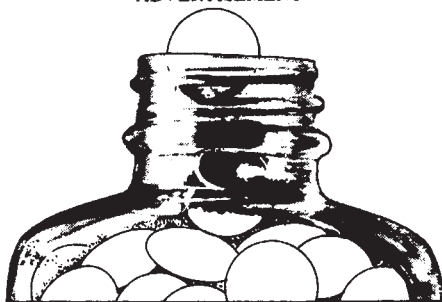
The argument throughout the book is lucid, ordered and informed. From the table of contents, one might suppose that much of it is mere exposition of different schools of thought. In fact, the author's treatment is more personal and more complex than the headings suggest. At each stage, ideas and events are placed in the context of his own view of the development process, which emerges from the volume as a whole rather than as an isolated statement. Despite its length, the book is not in the least discursive or repetitive, and at various points the treatment could with advantage have been made longer and more explicit. Little has something interesting and worth while to say on a wide range of issues and topics; even the footnotes, of which there are over 700, include a good many statements that bear pondering.

To my mind, the book's chief limitations are two, which are linked one to another. First, a lot of space is given to the ideas of academic economists. Not only does this make the argument less accessible to the general reader, but it looks odd when one bears in mind the fact — fully recognized by Little — that the influence of these ideas on actual policies and outcomes is uncertain, and probably rather slight. Second, not much is said about how and why government policies came to be what they were. Little is well aware that governments do not necessarily see themselves as concerned with development in his sense — that is, "a rise in the present value of average (weighted) consumption per head" (p.6). At various points he mentions in passing the influence of other motives, objectives and preoccupations. But like most economists, he takes the matter no further. Thus a central question in the interpretation of events is dealt with by setting it on one side.

Perhaps it might have been better if instead of writing one long book on economic development, Ian Little had produced two medium-sized ones, for rather different though overlapping audiences. None the less, this is a most impressive piece of work, which will be read with profit by those who reject the author's conception of development as well as those who share it. □

P.D. Henderson is a Professor in the Department of Political Economy at University College London.

ADVERTISEMENT



REGULATION AND RESTRAINT IN CONTEMPORARY MEDICINE IN THE UK AND USA

Edited by **Dr. Hugh L'Etang**,
Editor of The Practitioner
December 1982 £32.00 224pp
ISBN 0 333 34095 7

Published jointly by the Royal
Society of Medicine and
Macmillan Press Ltd.

Presenting the edited proceedings of an Anglo-American conference, held in October 1981, this book examines in detail the increasing regulatory pressures placed upon medical practice and research both from inside and outside of the profession. These pressures have led to a conflicting array of self-imposed as well as legally imposed constraints, and the contributors discuss here their desirability for both the patient and the profession.

Contents

The papers are divided into five sessions:

Session I

Medicine in a regulated world

Session II

Factors which can harm the patient

Session III

Regulation and Clinical Trials

Session IV

Constraints in General and

Consultant Practice

Session V

How much regulation is desirable?

Details of papers and contributors available on application.

Please send orders and requests for information to:

Felicity Parker, Product Manager
Scientific and Medical Division
The Macmillan Press Ltd.
Houndmills
Basingstoke, Hampshire RG21 2XS, UK.

SCIENTIFIC
& MEDICAL

M

MACMILLAN
PRESS LTD

Circle No.60 on Reader Service Card.

Blood and thunder

Tam Dalyell

Partial Progress: The Politics of Science and Technology.

By David Albury and Joseph Schwartz.
Pluto: 1982. Pp.215; £4.95. \$??.

Partial Progress is a polemic book. The central argument of the authors, David Albury and Joseph Schwartz, is that science and technology have been designed to serve particular purposes and reflect the interests of managers, employers and governments. Further, they seek to "demystify" the notion that the work of scientists and technologists has been inherently beneficial or morally neutral, and try to establish "rules to operate by when dealing with scientists". All this is stimulating reading, but I find myself provoked by some of the examples Albury and Schwartz have chosen and by the harshness of several of their judgements.

The authors say that, in the academic world, the budget cuts of the 1980s have been resolved entirely in favour of the new scientific managers, and assert that "the long eight-year apprenticeship starting with the three-year research studentship followed by a five-year post-doctoral fellowship" has led to no permanent job for almost one-third of the research work force. Assuredly, I agree that this is a most acute problem. Equally true is the contention that the place of the *unemployable* (author's italics) post-doctoral worker is taken by the next generation of "students". Albury and Schwartz, however, are intemperate in extending their argument: "The tenured professors continue to have a source of cheap labour for their labs and the whole revolving door is justified on the grounds that one needs 'fresh blood' to sustain a research programme". This last is a base accusation; it is simply not my experience in going around the institutions of higher education in Britain that the tortured souls who run departments — desperately worried about keeping a balanced age profile in their laboratories — are guilty of such a charge.

Albury and Schwartz attack the "fresh blood" argument for having served scientific managers and government administrators well, since, in justification of redundancies, they have been able to exploit the widely held belief that you have to be young to be productive in research. The opine that the level of political consciousness (and trade union organization) among the workforce was so low that the fresh blood argument passed unchallenged. According to the authors, the myth of the superior originality of young researchers began in the 1920s with the *Wunderkinder* — Pauli, Heisenberg and Dirac — involved in the creation of quantum mechanics and the dramatic new theory of the atom. The general situation,

however, is quite different. Elsevier's collection of Nobel Prize speeches has the relevant data: the average age of Nobel Prize winning physicists, *when they did their prize-winning work*, is 37 years old; in biology, it is 39 years old.

Rightly, in my view, the authors consider that the myth of youthful genius — and there is such a myth — obscures the fact that research is a collective process. True, some researchers were indeed very young when they did important work. Yet such youthful innovators were set to work on well-defined problems by those who were among the world's leading specialists in their field. Pauli, Heisenberg and Dirac all had the good fortune to work with Max Born at Gottingen, when he was in his late thirties, and an experienced member of the international community of atomic physics researchers. Later, when a teenager, I was to come to know Born and can easily understand how this marvellous old man could have been inspirational in his middle years. But I would like a better example than that of Born's relationship with Pauli, Heisenberg, and Dirac to convince me of Albury and Schwartz's conclusion that it is because of the myth of youthful genius, and despite changing working conditions and the threat of redundancy, that most scientists continue to see themselves as professionals.

Nor in my experience is it anything like the whole story to suggest that the response of scientists to changing conditions, and to the way that they are being treated is to say, "You cannot treat us like that. It is all very well for manual workers to be sacked at will, to be told how fast to work, what to do and so on, but we as scientists are different from that, we should not be treated like this". On the contrary, the reaction I have found about being forced into early retirement is one of, "Now we know what it is like for those who find themselves out of work, on the scrapheap, who were previously statistics to us". I do not think I am being unduly charitable. Rather I find that Albury and Schwartz have an excessively harsh view of the human qualities of their fellow scientists.

Finally, I cannot agree that it is because of their "training and class background" that scientists' attitudes towards their situation have not kept pace with their changed working conditions. But the authors are right to draw attention to the Association of University Teachers estimates that there are upwards of 2,500 scientists, all holders of doctorates, permanently out of work in Britain. This is truly a terrible indictment of the way in which the nation has contrived to manage its affairs. Despite exaggeration of their case, Albury and Schwartz deserve a wide audience. □

Tam Dalyell has been Member of Parliament for West Lothian since 1962. He was Shadow Spokesman on Science from 1980 to 1982, when he was dismissed over the Falklands issue.

Origin of Hawaiian tholeiite and alkalic basalt

Chu-Yung Chen* & Frederick A. Frey

Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Haleakala Volcano consists of a tholeiitic basalt shield overlain by alkalic lavas. As a function of decreasing age, La/Ce, Ba/La, Nb/La, Rb/Sr and $^{143}\text{Nd}/^{144}\text{Nd}$ increase but Sm/Nd and $^{87}\text{Sr}/^{86}\text{Sr}$ decrease. Inverse correlations between these isotopic ratios and their respective parent/daughter abundance ratios are consistent with mixing of melts derived from an enriched mantle composition with melts formed by <1% melting of a mid-ocean ridge basalt source.

BASALT is the most abundant volcanic rock produced within the Earth's mantle. Although several basalt types have been defined, basalt compositions define a continuum^{1,2}. Moreover, different basalt types often occur within a single volcanic structure. For example, Hawaiian volcanoes appear to be composed primarily of tholeiitic basalt, however, alkalic basalt dominates the last stages of volcanism³, and may precede tholeiites during early shield formation⁴. A major problem in understanding the origin of basalts is the relationship between associated tholeiitic and alkalic basalts. In particular, an understanding of the transition between tholeiitic and alkalic Hawaiian basalt would provide constraints on the mantle sources and processes involved in creating the volcanoes that form the linear Emperor Seamount and Hawaiian Ridge volcanic chains.

Early researchers studying the relation between Hawaiian tholeiitic and alkalic basalts suggested derivation of alkalic basalt from a tholeiitic parent by crystal fractionation^{5,6}, although experimental studies subsequently showed that this process could not occur at low pressures, thus these basalt types are not related by crustal processes. A more recent model based on experimental petrology⁷ and trace element geochemistry⁸⁻¹⁰ is that Hawaiian tholeiitic and alkalic basalt form from a compositionally homogeneous source with alkalic basalt representing magmas segregated at a greater depth and formed by a lower degree of melting than the associated tholeiitic magmas. In some cases, isotopic data, principally $^{87}\text{Sr}/^{86}\text{Sr}$, for associated alkalic and tholeiitic basalt suggest a common source^{11,12} whereas in other Hawaiian volcanoes systematic isotopic differences between alkalic and tholeiitic basalt^{11,13-15} preclude a common source. However, few studies include high-precision isotopic and trace-element analyses of well-characterized samples from documented stratigraphic sections of Hawaiian volcanoes; such studies are necessary for understanding the relationships between Hawaiian alkalic and tholeiitic basalt.

Haleakala Volcano forms East Maui and is comprised of a large tholeiitic shield (Honomanu series) which is almost completely overlain by younger alkalic basalt and hawaiite comprising the Kula and Hana series¹⁶; the latter is a post-erosional series formed dominantly of alkalic basalt with its most recent eruption in about 1790. To evaluate geochemical variations with eruption age, we studied Kula and Hana series samples from drill cores (up to ~300 m) obtained from the north-east slope of Haleakala¹⁶, and Honomanu series samples collected from the vicinity of Honomanu Bay. These samples encompass a major time fraction of the eruptive history of Haleakala, and our geochemical results indicate that the composition of the mantle source for Haleakala basalts changed systematically during evolution of the volcano. We interpret that the correlation of geochemical parameters with basalt age resulted from varying mixing proportions of components derived from two

Table 1 Chemical composition and isotopic ratios (Sr and Nd) of three typical Haleakala basalts

	Honomanu tholeiite C122	Kula alkalic basalt H65-4	Hana alkalic basalt H65-11
SiO ₂	48.05	44.04	44.51
TiO ₂	1.83	3.12	2.99
Al ₂ O ₃	9.89	14.28	14.14
Fe ₂ O ₃	13.44	15.38	15.43
MnO	0.17	0.20	0.20
MgO	16.31	8.23	8.89
CaO	7.77	11.04	10.61
Na ₂ O	1.67	2.94	2.86
K ₂ O	0.26	0.87	0.82
P ₂ O ₅	0.16	0.39	0.35
SUM	99.55	100.49	100.79
H ₂ O+	0.55	0.21	0.33
CO ₂	0.20	0.20	0.18
Rb	4.11	19.86	18.83
Sr	207	616	582
Ba	57	350	391
Sc	25.8	31.0	28.1
V	205	346	324
Cr	804	167	289
Co	76.9	61.7	63.9
Ni	703	109	139
Zn	115	110	109
Zr	103	200	166
Nb	7.2	33.3	31.1
Hf	2.64	4.87	4.12
Ta	0.5	2.1	2.1
Th	0.8	2.1	2.6
La	7.1	25.9	23.7
Ce	18.8	58.0	50.6
Nd	12.3	30.7	26.1
Sm	3.73	6.85	6.01
Eu	1.39	2.38	2.21
Tb	0.71	1.08	0.97
Ho	0.8	1.0	0.9
Yb	1.73	1.92	1.83
Lu	0.24	0.27	0.27
$^{87}\text{Sr}/^{86}\text{Sr}$	0.70387 ± 4	0.70334 ± 3	0.70314 ± 4
$^{143}\text{Nd}/^{144}\text{Nd}$	0.51291 ± 2	0.51302 ± 1	0.51310 ± 2

Trace elements, Ba, V, Ni, Zn, Zr and Nb, and major elements except Na₂O and K₂O determined by X-ray fluorescence at University of Massachusetts. K₂O and Sr determined by isotope dilution at MIT. REE, Sc, Cr, Co, Hf, Ta, Th and Na₂O determined by instrumental neutron activation at MIT. H₂O+ and CO₂ by CHN analyser (Woods Hole Oceanographic Institute). $^{87}\text{Sr}/^{86}\text{Sr}$ normalized to $^{87}\text{Sr}/^{86}\text{Sr} = 0.11940$ and relative to 0.70800 for E and A Standard. $^{143}\text{Nd}/^{144}\text{Nd}$ normalized to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$ and relative to 0.512646 for BCR-1. 2 σ errors refer to the last digit.

* Present address: Environmental Research Laboratory, University of Illinois, Urbana, Illinois 61801, USA.

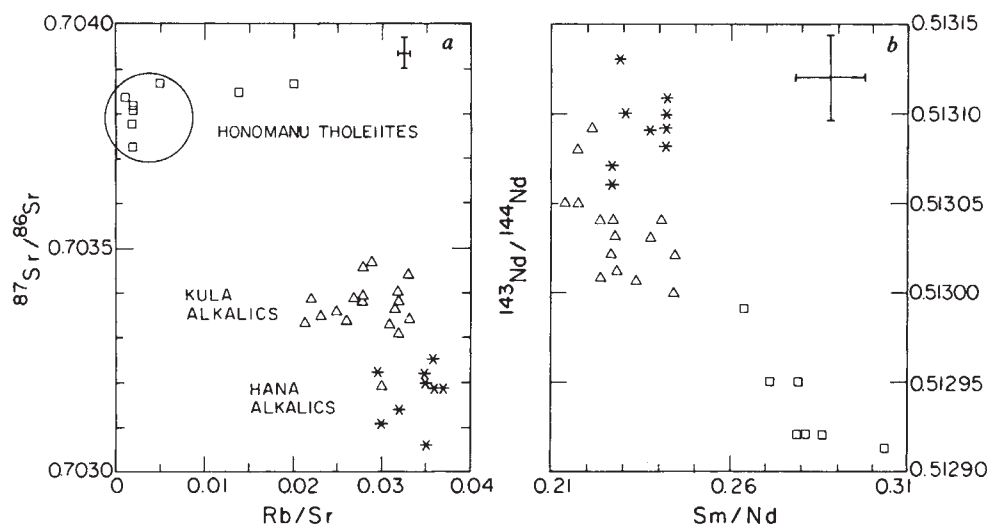


Fig. 1 *a*, $^{87}\text{Sr}/^{86}\text{Sr}$ ratio versus Rb/Sr abundance ratio. *b*, $^{143}\text{Nd}/^{144}\text{Nd}$ ratio versus Sm/Nd abundance ratio. Data are for Haleakala basalts: $\square$, the dominantly tholeiitic Honomanu series; $\triangle$, Kula alkalic series; *, post-erosional, Hana alkalic series. Tholeiites circled in left diagram contain olivine with altered rims, the extremely low Rb/Sr ratios are interpreted as a result of alteration. A similar explanation has been proposed for very low Rb/Sr in basalts from Kohala Volcano, Hawaii¹². Symbols in upper right indicate $\pm 2\sigma$ uncertainty.

compositionally distinct types of mantle, a mantle source suitable for generating normal mid-ocean ridge basalt (MORB) and a more enriched mantle source, having higher abundances of incompatible elements, lower $^{143}\text{Nd}/^{144}\text{Nd}$ and higher $^{87}\text{Sr}/^{86}\text{Sr}$ than a MORB source.

Results

We determined major and trace element abundances, $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ isotopic ratios in 11 samples from the dominantly tholeiitic Honomanu series, 24 samples from the alkalic Kula series and 10 samples from post-erosional, alkalic Hana series. The most striking feature of these data are the negative correlations between parent/daughter abundance ratios and radiogenic isotope abundances in the Rb–Sr and Sm–Nd systems (Fig. 1). These inverse trends are in marked contrast to the positive $^{87}\text{Sr}/^{86}\text{Sr}$ –Rb/Sr trends in some oceanic island basalt suites and the ‘mantle isochron’ concept proposed by Brooks *et al.*¹⁷. Moreover, ratios of incompatible elements, such as Nb/La, Ba/La and La/Ce, and isotopic ratios change systematically with age: the youngest alkalic series (Hana) has the highest Nb/La, Ba/La, La/Ce and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios and the lowest $^{87}\text{Sr}/^{86}\text{Sr}$ ratios whereas the older dominantly tholeiitic series (Honomanu) has the lowest abundances of highly incompatible elements, the lowest Nb/La, Ba/La, La/Ce and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios and the highest $^{87}\text{Sr}/^{86}\text{Sr}$ (Table 1). Because our objective is to infer source characteristics of these Haleakala basalts, we emphasize geochemical parameters that are not readily changed by fractional crystallization: that is, isotopic ratios and abundance ratios of highly incompatible elements such as La/Ce.

Interpretation

Although alkalic basalts typically have relatively high Rb/Sr and low Sm/Nd abundance ratios compared with tholeiitic basalts, they commonly have $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ isotopic ratios reflecting derivation from sources with very low Rb/Sr and high Sm/Nd (refs 18–20). Possible explanations for these results are that such alkalic basalts were derived by very low (<1%) degrees of melting so that parent/daughter abundance ratios were changed markedly during partial melting or that the mantle source of alkalic basalt experienced mantle metasomatism shortly before melting²¹. The geochemical trends of the tholeiitic and alkalic basalts at Haleakala (Fig. 1) enable us to propose a geodynamic model which may have general application to Hawaiian volcanism and possibly to other areas where tholeiitic and alkalic basalt have erupted.

If there is a genetic relationship between basalts which have an inverse correlation between radiogenic isotope abundance and corresponding parent/daughter abundance ratios (Fig. 1),

the most plausible explanation is a mixing process. In the mantle array of Sr and Nd isotopic ratios the youngest alkalic series (Hana) overlaps with the MORB field whereas the oldest tholeiitic series (Honomanu) lies closer to the estimated bulk Earth ratios (Fig. 2). Thus, the range in Sr and Nd isotope ratios in these Haleakala samples could be explained by mixing of two components; a component with isotopic ratios similar to MORB and a component with lower $^{143}\text{Nd}/^{144}\text{Nd}$ (<0.5129) and higher $^{87}\text{Sr}/^{86}\text{Sr}$ (>0.7039), the lowest and highest values, respectively, found in Haleakala tholeiites (Fig. 2).

However, a successful model for Haleakala basalts must also explain their systematic variation in trace-element abundance ratios. An important constraint on mixing models for Haleakala basalts is that the alkalic basalts which are MORB-like in their isotopic ratios are characterized by relatively high Rb/Sr, La/Ce and Nb/La and low Sm/Nd (Table 1). Melts formed by very low (<1%) degrees of melting of a MORB-source would have these geochemical characteristics: MORB isotopic ratios with high Rb/Sr, La/Ce, Nb/La and low Sm/Nd. Consequently,

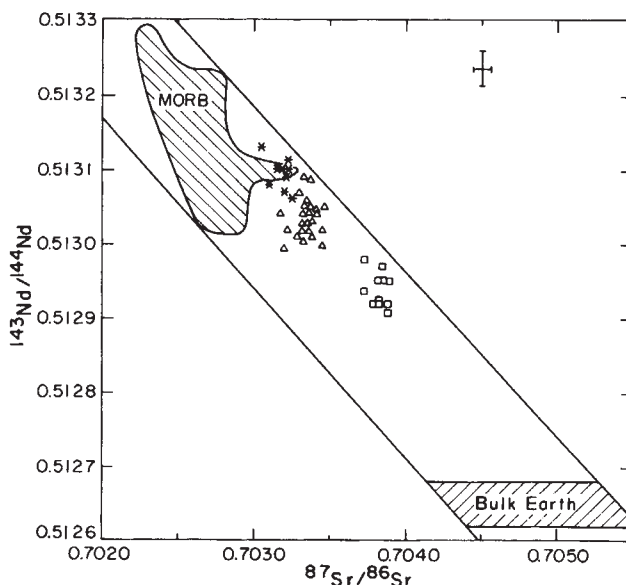


Fig. 2 $^{143}\text{Nd}/^{144}\text{Nd}$ versus $^{87}\text{Sr}/^{86}\text{Sr}$ showing field for mantle array defined by basalts. Shaded field^{23,24,34–36} indicates region for MORB which defines the high $^{143}\text{Nd}/^{144}\text{Nd}$ and low $^{87}\text{Sr}/^{86}\text{Sr}$ end of the mantle array. Haleakala basalts (symbols as in Fig. 1) lie between the MORB field and estimates for bulk-Earth values^{23–25,35,36}. Symbol in upper right indicates $\pm 2\sigma$ uncertainty.

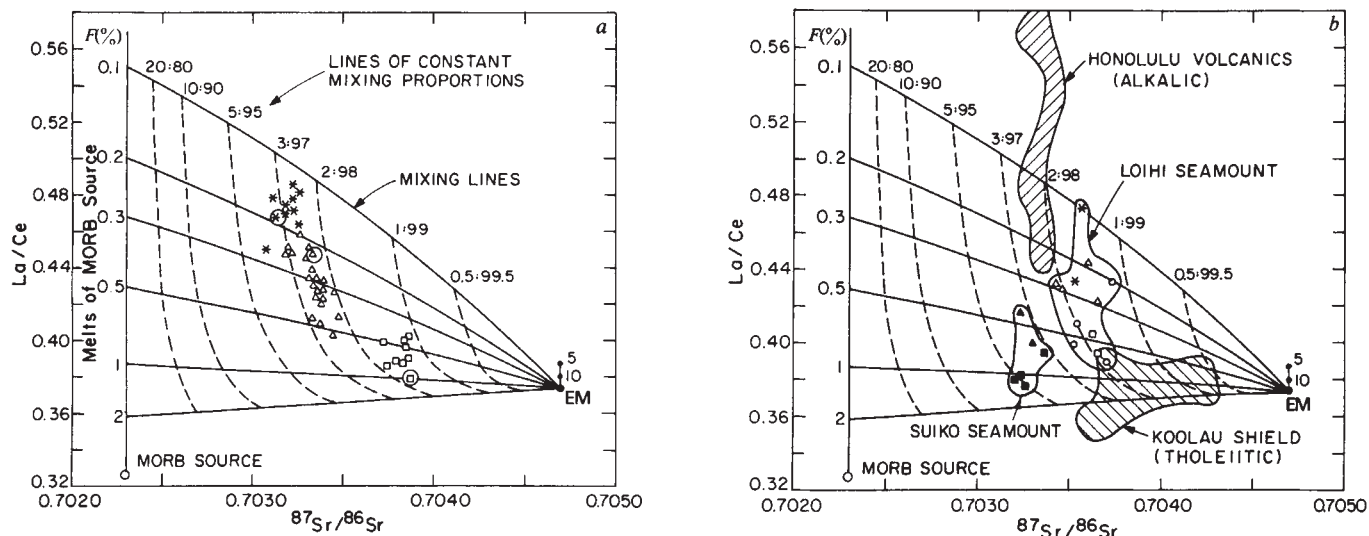


Fig. 3 La/Ce abundance ratio versus $^{87}\text{Sr}/^{86}\text{Sr}$ for Hawaiian basalts. **a**, Data points for Haleakala basalts. Honomanu series, primarily tholeiites ($\square$), Kula alkalic series (Δ), and Hana post-erosional alkalic series (*). The three samples used to test the model quantitatively (see Table 2 and Fig. 4) are circled. The vertical axis from the MORB source point indicates La/Ce abundance ratio as function of melting degree (F) for 0.1–2% melting of a MORB source containing 0.31 p.p.m. La, 0.95 p.p.m. Ce and 13.2 p.p.m. Sr. An equilibrium non-modal melting model was used with La and Ce partition coefficients from ref. 37 (set 1), and Sr partition coefficients: cpx/melt = 0.072, opx/melt = 0.0002, olivine/melt = 0.0005 and garnet/melt = 0.007. The initial mode of the MORB source was assumed to be olivine (60%), orthopyroxene (25%), clinopyroxene (10%) and garnet (5%) with the melt composed of olivine (10%), orthopyroxene (10%), clinopyroxene (40%) and garnet (40%). The solid lines with negative slope are mixing lines for mixtures of enriched mantle (EM) (La = 0.71 p.p.m., Ce = 1.90 p.p.m., Sr = 23.7 p.p.m.) with melts formed by low (<2%) degrees of melting of a MORB source. Mixing lines on such diagrams are typically curved³⁸, but on this diagram mixing lines are nearly straight because $(\text{Ce}/\text{Sr})_{\text{MORB source melt}}/(\text{Ce}/\text{Sr})_{\text{EM}}$ is ~ 1 (1.36–1.0 for 0.1–2% melting of the MORB source). The dashed, near vertical lines connect points with equal ratios of end member components: 5:95 indicates 5% of the melt from a MORB source and 95% EM. The vertical axis from EM indicates La/Ce abundance ratio in melts derived by 5 and 10% melting of the EM composition (assumed model parameters are same as for melting of the MORB source). Mixing lines for mixing of melts derived from a MORB source and EM would connect points along the vertical axis emanating from the MORB source with points along the vertical axis emanating from the EM. Detailed predictions of the model depend on the chosen partition coefficients, initial modes, source compositions and melt modes but the general applicability of the model does not depend on these parameters. Changes in these parameters will result in only small variations in degrees of melting and mixing proportions. **b**, Fields for other Hawaiian basalts. Honolulu volcanics are a post-erosional suite ranging from alkali olivine basalt to nepheline melillite³⁹. Their wide range in La/Ce at a uniform $^{87}\text{Sr}/^{86}\text{Sr}$ requires constant mixing proportions, but a wide range (0.05–0.25%) in melting degree of a MORB source. Field encloses data for 15 basalts (refs 13, 39 and M.F. Roden and F.A.F., unpublished data). Koolau shield tholeiites underlie the Honolulu volcanics and have distinctly higher and variable $^{87}\text{Sr}/^{86}\text{Sr}$. Compared with the alkalic Honolulu volcanics, the Koolau tholeiites require higher degrees of melting of a MORB source, but smaller amounts of this melt mixed with EM in variable proportions. Field encloses data for seven tholeiites (refs 13, 40 and M.F. Roden and F.A.F., unpublished data). Loihi Seamount consists of tholeiitic and alkalic basalts⁴; the latter have higher La/Ce and lower $^{87}\text{Sr}/^{86}\text{Sr}$ than the tholeiites^{15,41}. *, basanites; Δ , alkalic basalts; $\square$, tholeiites; $\circ$, translational basalts. Suiko Seamount in the Central Emperor Ridge consists of alkalic basalts (Δ) which overlie tholeiites ($\square$). Data from refs 11, 42 and 43. Note that these samples have the lowest $^{87}\text{Sr}/^{86}\text{Sr}$ ratios and required the largest amount of a component derived from a MORB source.

we evaluate a mixing model involving incipient melts from a MORB source mixing with components from an enriched mantle.

The enriched mantle component is not well defined, except that for the Haleakala data it must have $^{143}\text{Nd}/^{144}\text{Nd} < 0.5129$ and $^{87}\text{Sr}/^{86}\text{Sr} > 0.7039$ (these ratios are significantly lower and higher, respectively, than in a MORB source, Fig. 2). However, some Hawaiian tholeiites have even higher $^{87}\text{Sr}/^{86}\text{Sr}$; particularly Koolau tholeiites which range from 0.7036 to 0.7045 (refs 13, 22 and M.F. Roden and F.A.F. unpublished data). Consequently, for modelling we chose an enriched mantle composition with Sr and Nd isotopic ratios similar to bulk Earth estimates ($^{87}\text{Sr}/^{86}\text{Sr} = 0.7047$ (refs 22–24) and $^{143}\text{Nd}/^{144}\text{Nd} = 0.51265$ (refs 22–25), and incompatible element abundances similar to those estimated for primitive mantle (see captions to Figs 3 and 4). In what follows 'enriched mantle' designates this model composition, but this composition is not uniquely defined by the data.

The mixing model is illustrated in Fig. 3a. A key feature is that the incipient melts of the MORB source are not constrained to a fixed degree of melting. Also, the enriched mantle component could be a solid or partial melts derived from an enriched mantle. Thus, this model has two isotopically distinct end mem-

bers, but an infinite variety of end members with different trace-element abundances. For clarity, the mixing lines in Fig. 3 are drawn for mixing of enriched mantle with incipient melts derived from a MORB source but perhaps a physically more plausible model involves mixing of incipient melts from a MORB source with partial melts derived from an enriched mantle. Both models can explain the geochemical data for the Haleakala samples.

For example, the $^{87}\text{Sr}/^{86}\text{Sr}$ and La/Ce variations in these Haleakala samples can be derived from sources which are mixtures of 96–99% enriched mantle with 1–4% melt formed by 0.1–1% melting of a MORB source (Fig. 3a). To match the La concentrations in the tholeiitic and alkalic basalts ~ 12 to 16% melting of this mixed source is required (Table 2). As an example of the model involving mixing of two melts we find that the $^{87}\text{Sr}/^{86}\text{Sr}$ and La/Ce ratios and La content of tholeiite, C-122, could be formed by mixing 10% of a partial melt derived by 1% melting of a MORB source with 90% of a melt formed by 14% melting of an enriched mantle, and $^{87}\text{Sr}/^{86}\text{Sr}$, La/Ce and La content of alkalic basalt H65-4 could have resulted from mixing 22% of a partial melt derived by 0.25% melting of a MORB source with 78% of a melt formed by 9% melting of an enriched mantle (Table 2).

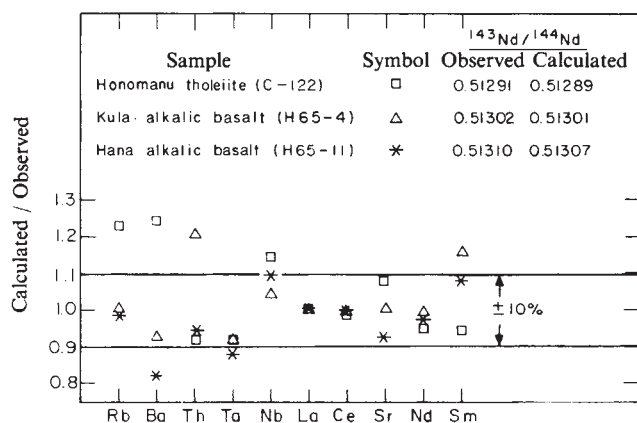


Fig. 4 Quantitative evaluation of the mixing end members and their proportions deduced from Fig. 3 (see Table 2) for the three representative samples listed in Table 1. Predicted Nd isotopic ratios are in excellent agreement with observed ratios. The vertical axis indicates the ratio of calculated to observed abundances for highly and moderately incompatible trace elements. Most of the calculated data are within $\pm 10\%$ of the observed abundances and all data points fit to better than $\pm 25\%$. To minimize complexities caused by fractional crystallization and uncertainties in partition coefficients, we modelled only highly and moderately incompatible elements in MgO-rich samples which appear to represent melts. The high MgO content (16.6%) and Mg/(Mg + Fe²⁺) ratio of the tholeiite are consistent with equilibrium olivine of Fo₈₉ to Fo₉₁; thus the trace-element abundances were not corrected for fractionation. However, the two alkalic basalts, H65-4 and H65-11, were corrected for 25% and 21% olivine loss. These corrections are first-order approximations and assume only olivine as a fractionating phase from a parental melt with 16.6% MgO. Calculated abundances were determined by using the mixing end members and proportions in Table 2. A non-modal equilibrium melting model⁴⁴, $C^e/C^o = 1/(D_o + F(1 - P))$, was used for melting of the various sources (MORB source, enriched mantle and mixed source). Initial mode and melting proportions were assumed to be the same in each case and are given in the legend to Fig. 3. The enriched mantle and the mixed source may also have contained minor and accessory phases such as phlogopite, amphibole and apatite, but we assumed that these phases were eliminated by the melting process. For La, Ce and Sr, the source abundances and partition coefficients are given in the legend to Fig. 3. Mineral/melt partition coefficients for Nd and Sm are from set 1 of ref. 37, for Rb and Ba they were assumed to be zero, and for Th, Nb and Ta a value of 0.001 was assumed. Initial abundances (C^o) in p.p.m. were assumed to be: C^o MORB source/ C^o enriched mantle: Rb, 0.1/0.73; Ba, 2.0/8.0; Th, 0.02/0.086; Ta, 0.016/0.049; Nb, 0.31/0.86; Nd, 0.86/1.2; Sm, 0.32/0.39.

These models and more specifically, the end-member components, their mixing proportions and the degrees of melting (Fig. 3a and Table 2), were formulated so that they satisfactorily account for the range of $^{87}\text{Sr}/^{86}\text{Sr}$ and La/Ce ratios and La abundances in Haleakala basalts. An important test of the models is to use the model parameters in Table 2 to infer $^{143}\text{Nd}/^{144}\text{Nd}$ and abundances of incompatible elements in Haleakala basalts. As shown in Fig. 4 incompatible element abundances predicted by the model agree (generally better than $\pm 10\%$) with the observed abundances for tholeiitic and alkalic basalt, and the model also accurately predicts the observed $^{143}\text{Nd}/^{144}\text{Nd}$ ratios. Moreover, derivation of Hawaiian basalts by mixing of components from this enriched mantle with incipient melts from a MORB source is also consistent with K_2O contents and isotopic ratios of Pb and He. Tatsumoto²⁶ concluded that Pb isotopic ratios in Hawaiian basalts are consistent with an asthenosphere plume-lithosphere (MORB-related) mixing model, and Kurz *et al.*²⁷ found that the $^3\text{He}/^4\text{He}$ - $^{87}\text{Sr}/^{86}\text{Sr}$ trend formed by Hawaiian basalts is consistent with mixing of a MORB reservoir with a more primitive

Table 2 Comparison of two mixing models for Haleakala basalts

	C122 tholeiite— Honomanu series	H65-4 alkalic basalt—Kula series	H65-11 alkalic basalt—Hana series
Mixing of melt from a MORB source with EM; data inferred from positions of these samples in Fig. 3a			
Weight ratio MORB source melt/EM	1.7/98.3	2.5/97.5	3.3/96.7
% Melting of MORB source	1	0.25	0.2
% Melting of mixed source (calculated to fit observed La abundance)	16	12	16
Mixing of melt from a MORB source with a melt derived from EM			
Weight ratio MORB source melt/EM melt	10/90	22/78	21/79
% Melting of MORB source	1	0.25	0.2
% Melting of EM	14	9	13

EM, enriched mantle.

source having high $^3\text{He}/^4\text{He}$ ratios. In addition, the mixing line required to explain the $^3\text{He}/^4\text{He}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ variations in Hawaiian basalts²⁷ requires $(\text{He}/\text{Sr})_{\text{MORB reservoir}}/(\text{He}/\text{Sr})_{\text{primitive mantle}} \gg 1$; such values are only likely if the MORB component is an incipient melt.

We also evaluated models involving a MORB source mixing with melts formed by small degrees of melting of an enriched source. Although such a model can account for the La/Ce and $^{87}\text{Sr}/^{86}\text{Sr}$ variations in Haleakala lavas, it cannot satisfactorily explain abundances of other incompatible elements and $^{143}\text{Nd}/^{144}\text{Nd}$.

Model for evolution of Hawaiian volcanism

We suggest that the model developed in Fig. 3 and tested in Fig. 4 has general applicability for Hawaiian volcanism. For example, Loihi Seamount, the most recent manifestation of the Hawaiian hotspot, contains tholeiites with higher $^{87}\text{Sr}/^{86}\text{Sr}$ and lower $^{143}\text{Nd}/^{144}\text{Nd}$ than the associated alkalic basalts^{14,15}. Similarly, alkalic Honolulu Series basalts on Oahu have lower $^{87}\text{Sr}/^{86}\text{Sr}$ and higher $^{143}\text{Nd}/^{144}\text{Nd}$ than the underlying Koolau shield tholeiites^{13,19,22} (M. F. Roden *et al.* unpublished data). Figure 3b illustrates the nature and amounts of end members required to form basalts from Loihi and Suiko Seamounts and the Koolau shield with its overlying alkalic Honolulu Series.

The correlation of geochemical parameters with age leads us to propose a geodynamic model²⁸ for Hawaiian volcanism. Because of isotopic evidence (Sr, Nd and He) for an enriched mantle component in Hawaiian volcanics we assume that Hawaiian volcanism is initiated by rise and decompression melting of an enriched mantle plume from the lower mantle. Heat is transferred from this upwelling plume to its wall rocks which are presumed to be depleted mantle capable of generating MORB. Eventually, these wall rocks are heated to their solidus and incipient melts (0.05–2% melting) are formed which mix in varying proportion with melt derived from the enriched mantle plume to form the spectrum of basalts found in Hawaiian volcanoes such as Haleakala.

The initial eruptives of a Hawaiian volcano, characterized by the alkalic Loihi basalts, contain $\sim 1.5\%$ of a melt derived from the wall rock (MORB source) by very low (0.1–0.3%) degrees of melting (Fig. 3b). These low melting degrees of the wall rocks may characterize the onset of volcanism because the wall

rocks must initially be heated to their solidus temperature by the ascending plume.

With continued volcanism, the volcano evolves into the major, shield construction phase of Hawaiian volcanism (Honomanu series of Haleakala and Koolau shield) and the degrees of partial melting in the wall rocks increase to 0.5–2% (Fig. 3). Basalts generated at this stage are tholeiitic, contain a higher proportion of enriched mantle (Fig. 3) and consequently have higher $^3\text{He}/^4\text{He}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ but lower $^{143}\text{Nd}/^{144}\text{Nd}$ ratios than the alkalic basalts.

Eventually, the volcano migrates away from the hotspot and the supply of enriched mantle and heat are diminished. As a result, the eruption frequency decreases, alkalic basalts become dominant, degrees of partial melting in the wall-rocks decrease and the relative contribution of enriched mantle is decreased (Table 2). Basalts erupted at this stage are alkalic basalts such as the Kula Series at Haleakala. Finally, during the final stages of volcanism the melts formed by very low (0.05–0.2%) degrees of melting of a MORB source become increasingly important in determining abundance ratios, such as La/Ce, Ba/La and Nb/La, and isotopic ratios in the mixed magma. This stage is represented by the post-erosional Hana Series of Haleakala and the Honolulu Series on Oahu (Table 2 and Fig. 3).

Tholeiites from Suiko Seamount, central Emperor Ridge, have the lowest $^{87}\text{Sr}/^{86}\text{Sr}$ observed in Hawaiian tholeiites¹¹; consequently, they contain the largest amount of melt derived from a MORB source (Fig. 3b). Lanphere *et al.*¹¹ note that Suiko Seamount formed on younger oceanic crust than currently active Hawaiian volcanoes. Possibly, the thermal regime beneath younger crust enables a larger proportion of MORB source melt to mix with melts from the upwelling enriched mantle. The alkalic and tholeiitic basalts from Suiko Seamount plotted in Fig. 3b overlap in $^{87}\text{Sr}/^{86}\text{Sr}$ but Sr isotopic data for a group of Suiko Seamount lavas show that the mean for five tholeiitic samples (0.70324 ± 7) is slightly higher than the mean for four alkalic samples (0.70313 ± 15)¹¹.

Undoubtedly, the petrogenetic processes leading to Hawaiian volcanism are considerably more complex than the model outlined; however, the model adequately explains the major systematic variations in Sr, Nd, Pb and He isotopic ratios and abundances of highly incompatible elements. Moreover, the correlation of geochemical variations with eruption age leads to a plausible dynamic model for Hawaiian volcanism. An important implication of the model is that the large size, $20\text{--}30 \times 10^3 \text{ km}^3$, of many Hawaiian volcanoes²⁹ and the degrees of melting that we infer for the mixing components (Table 2) require that the entire lower lithosphere and much of the asthenosphere beneath a Hawaiian volcano must be involved in its creation. Another inference is that Hawaiian tholeiitic magmas segregated from their source at a greater depth than Hawaiian alkalic magmas.

General applicability of the mixing model

Wasserburg and De Paolo³⁰ proposed an Earth structure with an 'essentially undifferentiated' lower mantle overlain by a depleted upper mantle, and they invoked mixing between melts derived from lower mantle diapirs and oceanic upper mantle to account for the Sr and Nd isotopic characteristics of island basalts. Such an Earth structure is consistent with our model; however, the Haleakala data require considerably more complex mixing processes. Moreover, we can define the general geochemical characteristics of the mixing end members.

However, we emphasize that the Hawaiian geochemical data do not rigorously constrain the composition of the enriched component. In Figs 3 and 4, the composition of the enriched mantle corresponds to values proposed for primitive mantle, but we have also evaluated other compositions for the enriched component such as recently enriched mantle (higher incompatible element abundances), ancient enriched mantle (higher $^{87}\text{Sr}/^{86}\text{Sr}$ and incompatible element abundances and lower $^{143}\text{Nd}/^{144}\text{Nd}$) and recently depleted mantle (lower incompatible element abundances). Although the fit of these models to the data is poorer than the models in Figs 3 and 4, we conclude that the isotopic and incompatible element abundances of the enriched end member cannot be confidently constrained to the specific values used in Figs 3 and 4.

Despite this ambiguity in defining the enriched mantle composition, the proposed mixing model may be appropriate for other areas where basalts have an inverse correlation between Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$, for example, Nunivak Island¹⁸ on the Bering Sea shelf (also unpublished theses of R. K. Mark 1971 and M. F. Roden 1982); the Cameroon volcanic line³¹ which crosses a continental margin (J. G. Fitton, personal communication) and continental basalts such as the Afar volcanics, Ethiopia³² and eastern China³³. Furthermore, a positive correlation between Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$ results from the model when the degree of melting of the MORB source exceeds 1–2% (Fig. 3).

Consequently, the model proposed here may be applicable to several other oceanic and continental basalt suites, although it is likely that the compositions of the MORB source and enriched mantle are variable.

We thank D. Clague, M. Garcia, T. Grove, S. Hart, M. Roden, N. Shimizu and A. Zindler for constructive critical comments on the manuscript; the NSF for grant EAR-7823423 which supported this research, S. R. Hart for use of mass spectrometry facilities, G. Thompson for use of CHN analyser, and M. Rhodes for use of X-ray fluorescence equipment, also M. Garcia who initiated our interest in Haleakala and provided the samples for study, and M. Roden for discussions of Hawaiian volcanism and use of unpublished isotopic data for the Koolau and Honolulu volcanic series.

Received 20 September 1982; accepted 21 February 1983.

1. Wilkinson, J. F. G. *Basalts* Vol. 1, 163–214 (Interscience, New York, 1967).
2. Yoder, H. S. & Tilley, C. D. *J. Petrol.* **3**, 342–532 (1962).
3. Macdonald, G. A. & Katsura, T. *J. Petrol.* **5**, 82–133 (1964).
4. Moore, J. G., Clague, D. A. & Normark, W. R. *Geology* **10**, 88–92 (1982).
5. Powers, H. A. *Am. J. Sci.* **30**, 57–71 (1935).
6. Green, D. H. & Ringwood, A. E. *Contr. Miner. Petrol.* **15**, 103–190 (1967).
7. Presnall, D. C., Jackson, E. D. & Dixon, S. A. *Abstr. Vol. Hawaii Symp. on Intraplate Volcanism and Submarine Volcanism* **96** (1979).
8. Gast, P. W. *Geochim. cosmochim. Acta* **32**, 1057–1086 (1968).
9. Schilling, J.-G. & Winchester, J. W. *Contr. Miner. Petrol.* **23**, 27–37 (1969).
10. Feigenson, M. D. & Spera, F. J. *Geology* **9**, 531–533 (1981).
11. Lanphere, M. A., Dalrymple, G. B. & Clague, D. A. *Init. Rep. DSDP Leg 55*, 695–706 (1980).
12. Feigenson, M. D., Hofmann, A. W. & Spera, F. J. *Contr. Miner. Petrol.* (submitted).
13. Lanphere, M. A. & Dalrymple, G. B. *Am. J. Sci.* **280A**, 736–751 (1980).
14. Staudigel, H., Hart, S. R., Zindler, A., Lanphere, M., Chen, C.-Y. & Clague, D. *EOS* **62**, 1083 (1981).
15. Lanphere, M. *Earth planet. Sci. Lett.* (submitted).
16. Stearns, H. T. & Macdonald, G. A. *Hawaii Div. hydrol. Bull.* **7**, 344 (1942).
17. Brooks, C., Hart, S. R., Hofmann, A. & James, D. E. *Earth planet. Sci. Lett.* **32**, 51–61 (1976).
18. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **46**, 323–334 (1980).
19. Roden, M. F., Frey, F. A. & Hart, S. R. *EOS* **62**, 423 (1981).
20. Allègre, C. J., Dupre, B., Lambret, B. & Richard, P. *Earth planet. Sci. Lett.* **52**, 85–92, (1981).

21. Menzies, M. & Murthy, V. R. *Am. J. Sci.* **280**, 622–638 (1980).
22. De Paolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 743–746 (1976).
23. Allègre, C. J., Ben Othman, D., Polve, M. & Richard, P. *Phys. Earth planet. Inter.* **19**, 293–306 (1979).
24. O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13–22 (1977).
25. Jacobsen, S. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **50**, 139–155 (1980).
26. Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63–87 (1978).
27. Kurz, M. D., Jenkins, W. J. & Hart, S. R. *Nature* **297**, 43–47 (1982).
28. Allègre, C. J. *Tectonophysics* **81**, 109–132 (1982).
29. Bargar, K. E. & Jackson, E. D. *J. Res. U.S. geol. Surv.* **2**, 545–550 (1974).
30. Wasserburg, G. J. & DePaolo, D. J. *J. Proc. natn. Acad. Sci. U.S.A.* **76**, 3954–3958 (1979).
31. Dunlop, H. M. & Fitton, J. G. *Contr. Miner. Petrol.* **71**, 125–131 (1979).
32. Barberi, F., Civetta, L. & Varet, J. *Earth planet. Sci. Lett.* **50**, 247–259, (1980).
33. Zhou, X. & Armstrong, R. L. *Earth planet. Sci. Lett.* **58**, 301–329 (1982).
34. White, W. M. *Yb Carnegie Inst. Wash.* **78**, 325–331 (1979).
35. Cohen, R. S., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* **283**, 149–153 (1980).
36. Dupre, B. & Allègre, C. J. *Nature* **286**, 17–21 (1980).
37. Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463–513 (1978).
38. Langmuir, C. H., Vocke, R. D., Hanson, G. N. & Hart, S. R. *Earth planet. Sci. Lett.* **37**, 380–392 (1978).
39. Clague, D. A. & Frey, F. A. *J. Petrol.* **23**, 447–504 (1982).
40. Leeman, W. P., Budahn, J. R., Gerlach, D. C., Smith, D. R. & Powell, B. N. *Am. J. Sci.* **280**, 794–819 (1980).
41. Frey, F. A. & Clague, D. A. *Earth planet. Sci. Lett.* (submitted).
42. Clague, D. A. & Frey, F. A. *Init. Rep. DSDP leg 55*, 559–569 (1980).
43. Bence, A. E., Taylor, S. R. & Fisk, M. *Init. Rep. DSDP Leg 55*, 599–605 (1980).
44. Shaw, D. M. *Geochim. cosmochim. Acta* **34**, 237–243 (1970).

Mechanism of calcium channel blockade by verapamil, D600, diltiazem and nitrendipine in single dialysed heart cells

K. S. Lee* & R. W. Tsien†

Department of Physiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA

Organic inhibitors of calcium influx prevent outward as well as inward current through cardiac calcium channels but do not slow current activation. Although block is antagonized by raising external calcium or barium concentrations, the competitive effect of permeant cations does not occur at the same cation binding site at which inorganic blockers act. Organic drugs show varying degrees of use-dependent block, due in part to blockade of open channels. Nitrendipine blockade of calcium currents requires doses >100-fold higher than expected from radioligand binding to isolated membranes.

In many excitable cells, the entry of calcium ions can be inhibited by verapamil, D600, diltiazem and dihydropyridine compounds such as nifedipine or nitrendipine. These drugs are often grouped together as 'Ca antagonists' or 'Ca channel blockers'^{1,2,3}, although they represent quite diverse classes of organic compound^{3,4}. As therapeutic agents, they have proven effective in the management of cardiac arrhythmias and coronary disease (see, for example, refs 4,5). As ligands, they bind strongly to membranes from excitable cells⁶⁻⁹, and will no doubt be used in future attempts to count, purify and characterize Ca channels.

Although the organic Ca current inhibitors have been widely studied¹⁻⁵, several fundamental questions remain controversial or unanswered^{4,5}. (1) Do the organic compounds merely inhibit Ca influx, by limiting the supply of inwardly permeating ions, or do they prevent outward as well as inward current through Ca channels, as the label 'Ca channel blocker' implies? (2) Do the drugs slow the activation of Ca channels? (3) Do the inhibitory effects show competition with extracellular Ca, as the term 'Ca antagonist' requires? (4) Does competition occur at the same divalent cation binding site where inorganic ions such as Mn^{2+} , Co^{2+} or Cd^{2+} block Ca channels? (5) How do the various organic compounds differ in the way their blocking effects depend on prior membrane depolarization? (6) Does blockade of Ca channel current in functional cells parallel radioligand binding to isolated membranes?

These questions seemed well suited for analysis using the suction-pipette technique in single heart cells¹⁰⁻¹². Here we report systematic comparisons between the blocking effects of the three major classes of organic Ca antagonist, represented by verapamil or D600, diltiazem and nitrendipine, as well as comparisons between organic and inorganic blockers. The dihydropyridine compound nitrendipine was of particular interest: although its tritiated form has been used widely in binding experiments⁶⁻⁹, a systematic study of its electrophysiological action has not been reported.

Single isolated heart cells were obtained from ventricles of adult guinea pig hearts by enzymatic dissociation. A suction pipette (tip diameter 5–10 μm) provided access to the inside of the cell for voltage clamp and internal dialysis¹⁰⁻¹². The internal and external solutions contained Cs^+ in place of Na^+ or K^+ except in Fig. 1. Currents through sodium channels were eliminated by holding the membrane potential near -50 or -40 mV, and by the absence of Na or by the presence of tetrodotoxin

in the external solution. Depolarizing pulses were delivered at 3 or 6 per min. All experiments were carried out at room temperature (20–22 °C).

Blockade of outward current through Ca channels

When heart cells are strongly depolarized, their Ca channels can carry outward current^{12,13} due to the efflux of K^+ or Cs^+ . The outward current shows the kinetics and the Cd^{2+} -sensitivity expected for genuine Ca channel current¹². Here, we use recordings of such outward Ca channel current to help settle a long-standing debate about the mechanism of action of verapamil and other Ca current inhibitors. If verapamil reduces Ca influx by depleting a pool of extracellular charge carriers¹, one would expect outward movement of K^+ through Ca channels to be unaffected or perhaps increased. If, on the other hand, the drug acts through channel blockade, it should reduce current flow in either direction. As Fig. 1a illustrates, verapamil strongly inhibits outward as well as inward current through Ca channels. The same is true for diltiazem (Fig. 1b), nitrendipine (Fig. 1c) and verapamil's methoxy-derivative D600 (ref. 12). Although not illustrated here, the blockade of outward current through Ca channels was also seen when the outward current was carried by Cs^+ ions, either in the absence of permeant divalent cations or in the presence of external Ca^{2+} or Ba^{2+} . These results provide experimental justification for calling the organic drugs Ca channel blockers.

At the reversal potential (E_{rev}), where time-dependent current through the Ca channel vanishes, records with and without drug can be superimposed almost perfectly. This was a consistent finding for verapamil, diltiazem, nitrendipine, and also D600¹². In each case, the global current can be accounted for by Ca channel current in parallel with a time-independent, drug-insensitive leak current. Because other ionic currents seem virtually absent in these conditions, analysis of drug effects is simpler than in other preparations where Ca channel and K channel currents overlap (see ref. 14).

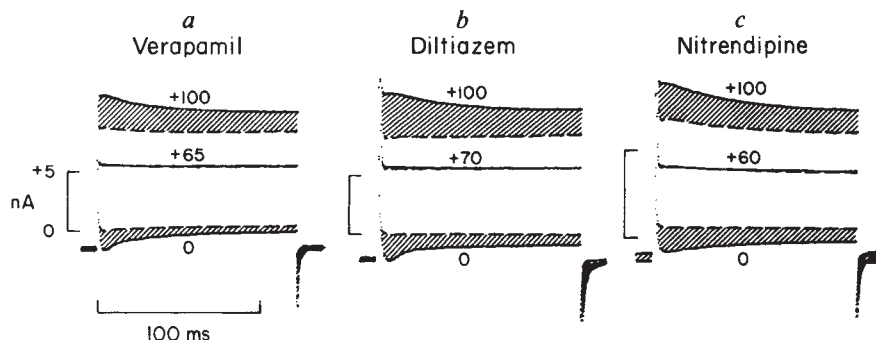
Activation kinetics

Do the organic drugs merely reduce the magnitude of Ca channel current, or do they also slow activation gating? In earlier sucrose gap experiments in cat ventricular muscle, Nawrath *et al.*¹⁵ found that D600 can greatly lengthen time-to-peak slow inward current. They suggested that the drug not only blocks channels, but also slows activation gating in unblocked channels by more than 10-fold. Such a dual effect would be very important if genuine, but the evidence from

* Present address: University Laboratory of Physiology, Oxford University, Parks Road, Oxford OX1 3PT, UK.

† To whom reprint requests should be addressed.

Fig. 1 Blockade of inward and outward current through Ca channels by 1 μM verapamil (a), 5 μM diltiazem (b) and 5 μM nitrendipine (c). Total membrane current records during depolarizing clamp pulses from a holding potential of -40 mV to the voltage levels indicated. Each panel superimposes records taken before drug (continuous traces) with those taken after about 2 min drug exposure (broken traces). Shaded areas denote inhibitor-sensitive outward current (at $+100$ mV) and inhibitor-sensitive inward current (at 0 mV). At the membrane potential where time-dependent current inverts (E_{rev}), near $+60$ or $+70$ mV, records from control and drug runs coincide almost perfectly. This is consistent with drug blockade of Ca channels without any effect on time-independent leak current. At $+100$ mV, some decaying outward current remained after exposure to 5 μM nitrendipine in four out of five cells, although not always as much as in c. Vertical calibration bars go from zero current up to $+5$ nA. External solution was 1 mM Ca Tyrode (137 mM NaCl, 11.9 mM NaHCO_3 , 0.42 NaH_2PO_4 , 5.4 KCl, 1.05 MgCl_2 , 1.0 CaCl_2 , 5.5 glucose and 10 μM tetrodotoxin, pH 7.4). Internal dialysate contained 151 mM K^+ , rather than 151 mM Cs^+ as described in Fig. 2 legend. Cells 140J, 141A and 140D in a–c, respectively.



other studies is contradictory (for example, compare refs 16–18 with ref. 19). We looked carefully for changes in activation kinetics, taking advantage of the relatively rapidly settling voltage clamp in single dialysed cells, but found no convincing evidence for any slowing. The time course of activation remained essentially unaffected during partial block of Ca current by D600 (Fig. 2a).

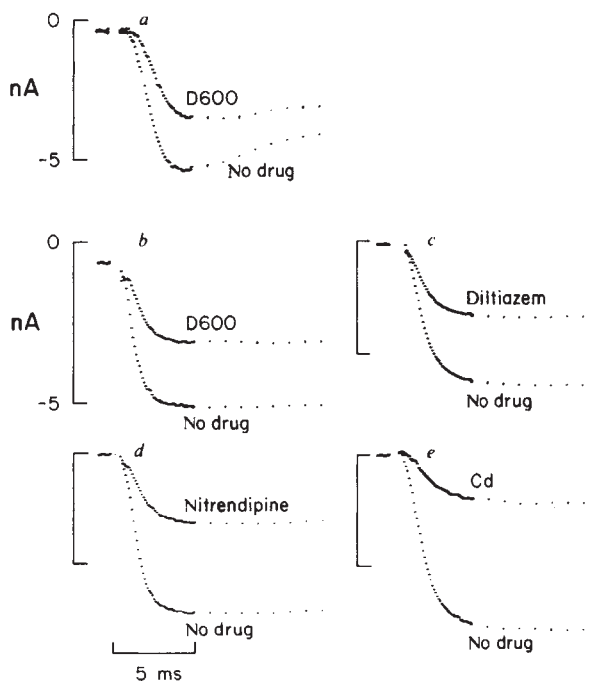
The current traces in Fig. 2a show a relatively rapid phase of inactivation that is characteristic of Ca channels with external Ca^{2+} as the current carrier^{12,14,17,18}. Inactivation is sufficiently rapid to overlap significantly with the activation process. To simplify the analysis of Ca channel activation, we replaced Ca^{2+} with Ba^{2+} , capitalizing on the slow inactivation of the Ba current. As Fig. 2b illustrates, the results support the conclusion that D600 leaves the activation kinetics unchanged. Figure 2c, d show similar behaviour during partial block of Ba currents by diltiazem and nitrendipine. In this respect, the organic Ca channel inhibitors resemble the inorganic ion Cd^{2+} (Fig. 2e).

Fig. 2 Activation of Ca channel current in the absence and presence of organic and inorganic inhibitors. Each panel superimposes a current record in the presence of drug (as labelled) with a corresponding record from a prior control run in the absence of drug ('no drug'). Inhibitor effects have reached steady-state in b (1 μM D600), c (5 μM diltiazem) and e (5 μM CdCl_2), but not in a (0.3 μM D600) or d (0.5 μM nitrendipine). Membrane current signals were obtained following digital subtraction of linear capacity and leak currents (determined with -10 mV test pulses). Holding potential was -40 mV. Test potential was $+20$ mV in a, c and d, $+40$ mV in b and $+10$ mV in e. Internal dialysate was standard 0-K, 151 mM-Cs solution containing 75.5 mM Cs-aspartate, 29 mM Cs_2HPO_4 , 17.5 mM CsH_2PO_4 , 5.5 mM glucose (pH 7.0). External solution was (mM) 0-Na, 0-K solution containing 85 Cs-aspartate, 2.85 MgCl_2 , 146 sucrose, 5.5 glucose (pH 7.4), plus a variable amount of CaCl_2 or BaCl_2 : 10 mM CaCl_2 in a, 20 mM BaCl_2 in b and 10 mM BaCl_2 in c–e. Vertical calibration bars from 0 to -5 nA in each panel. Current signals were filtered at 3 kHz. Sample rate was 10 kHz initially, then 1 kHz starting 5 ms after the onset of the voltage step. Data points are not shown for the first 300–500 μs after the onset of the voltage pulse while the membrane potential settled and the total current signal came into the range of the analog-to-digital converter. Cells 139A, 98D, 142G, 141C and 113E in a–e, respectively. We have considered two possible explanations for the slowed onset of inward current previously reported for D600 in cat ventricular muscle¹⁵. (1) The behaviour of the Ca channels might be obscured by transient outward current through K channels; 'slowed activation' would reflect the decay of outward current unmasked by calcium channel block. Transient outward current may be sizeable in cat papillary muscle^{14,15}. It was eliminated in our guinea pig cells by the -40 mV holding potential and by replacement of intracellular K with Cs. (2) The apparent slowing might arise from series resistance error in the single sucrose gap experiments¹⁵. In the present study, series resistance errors were reduced by the use of single cells and series resistance compensation, so that any resulting changes in time course were small. We also looked for changes in the development of outward Ca channel current because the effect of residual series resistance error would be reversed. Here, as in the case of inward Ca channel currents, the time course of activation remained essentially unchanged as the conductance decreased.

Calcium antagonism

Is blockade by the organic drugs antagonized by extracellular Ca^{2+} ? In their original work, Fleckenstein and colleagues proposed that verapamil and D600 compete with Ca for a common site^{1,20}. However, the restorative effects of elevated external Ca (Ca_o) in their experiments could be attributed to increased Ca influx through a constant fraction of unblocked channels. More recent experimental evidence on this point is contradictory. Payet *et al.*²¹ described a noncompetitive interaction between Ca^{2+} and verapamil or D600 in rat ventricular muscle, while Akaike *et al.*²² reported that diltiazem blockade of Ca current in snail neurones could be overcome by raising the extracellular Ca in accordance with simple competition.

Figure 3 compares the inhibitory effects of various blockers on the Ca current in the presence of 3 mM and 30 mM Ca. D600, diltiazem and nitrendipine show qualitatively similar behaviour. In each case, elevation of Ca_o significantly increases



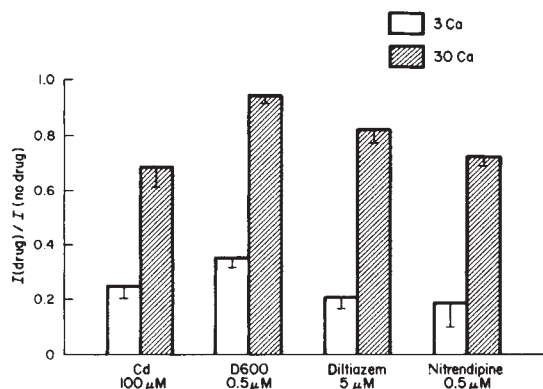


Fig. 3 Effect of external Ca on the degree of blockade. The ordinate plots the magnitude of the peak Ca current in the presence of inhibitor, as a fraction of peak Ca current during the preceding control run. Peak currents were measured relative to baseline after subtraction of the linear capacity transient and leak current. Each pair of histogram bars shows how the percentage of unblocked channels varies with the external Ca concentration for a fixed dose of inhibitor. Ca current was elicited by depolarizing pulses applied once every 10 s, from a holding potential of -40 mV up to the inward maximum of the current voltage relation ($+10$ mV in 3 mM Ca, and $+30$ mV in 30 mM Ca). Additional experiments (not illustrated) showed that the inhibitors merely scaled the current-voltage relationship without significantly altering its shape. External solutions were standard 0-Na, 0-K solution with either 3 or 30 mM CaCl_2 . Histogram bars and error brackets indicate mean and s.e.m. for three to seven observations.

the percentage of unblocked Ca channels. The competition with Ca^{2+} ions supports the common practice of designating all three kinds of organic compound as Ca antagonists.

Contrast between organic and inorganic blockers

In competing with Ca_o , the organic drugs resemble the inorganic blocker Cd^{2+} (Fig. 3). Do all of these agents act at a common site? Akaike *et al.*²² proposed that inorganic and organic blockers compete with Ca^{2+} for the same site within the channel, the metal cation coordination site described by Hagiwara *et al.* (see ref. 23 for review). We tested this hypothesis by comparing the responsiveness of Ca and Ba currents to the various blockers. Ba^{2+} is known to bind less strongly than Ca^{2+} to the metal ion coordination site²³; therefore, agents which act at that site should be more effective in reducing Ba currents than

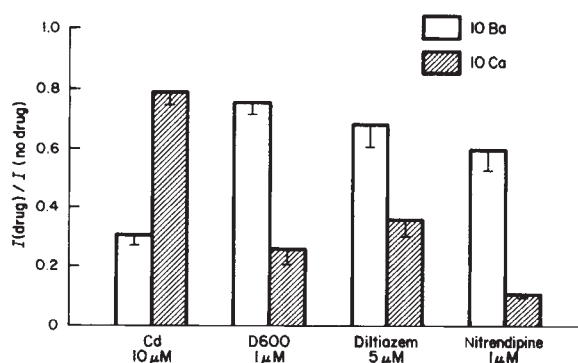


Fig. 4 Comparison between inhibitor block of Ca channel currents carried by Ca^{2+} or Ba^{2+} . Experimental procedure and graphical presentation are as in Fig. 3. Test depolarizations from -40 mV up to a level near the inward maximum of the peak current voltage relation ($+20$ mV in 10 mM Ca, $+10$ mV in 10 mM Ba). Bars and error brackets indicate mean and s.e.m. for three or four observations.

Ca currents. As expected, Cd^{2+} acted in just this way (Fig. 4, left). Surprisingly, however, D600, diltiazem and nitrendipine showed the opposite behaviour: each of these drugs was considerably less effective in blocking Ba currents than in blocking Ca currents. Thus, antagonism between permeant cations and organic blockers cannot be explained by simple competition at the coordination site of Hagiwara, but may involve Ca^{2+} or Ba^{2+} binding to another site with different cation preference. A cation binding site close to the inside of the membrane would be well positioned to influence the binding of organic drug molecules that approach the channel from the cytoplasmic side¹⁹.

Use-dependence

Do the structurally diverse organic compounds differ in their effects on channels? Accentuation of drug inhibition by repetitive membrane depolarization ('use-dependent' or 'frequency-dependent' block) is a prime area of contrast, judging by previous experiments²⁴⁻²⁹. Figure 5 compares the actions of D600, diltiazem and nitrendipine, using a protocol²⁹ similar to one used in studying local anaesthetic blockade of Na channels³⁰. Calcium currents were measured in the absence of drug with test depolarizations once every ~ 20 s. Depolarizations were then discontinued, drug was applied, and after a 3-min quiescent period, depolarizing test pulses were resumed. This procedure

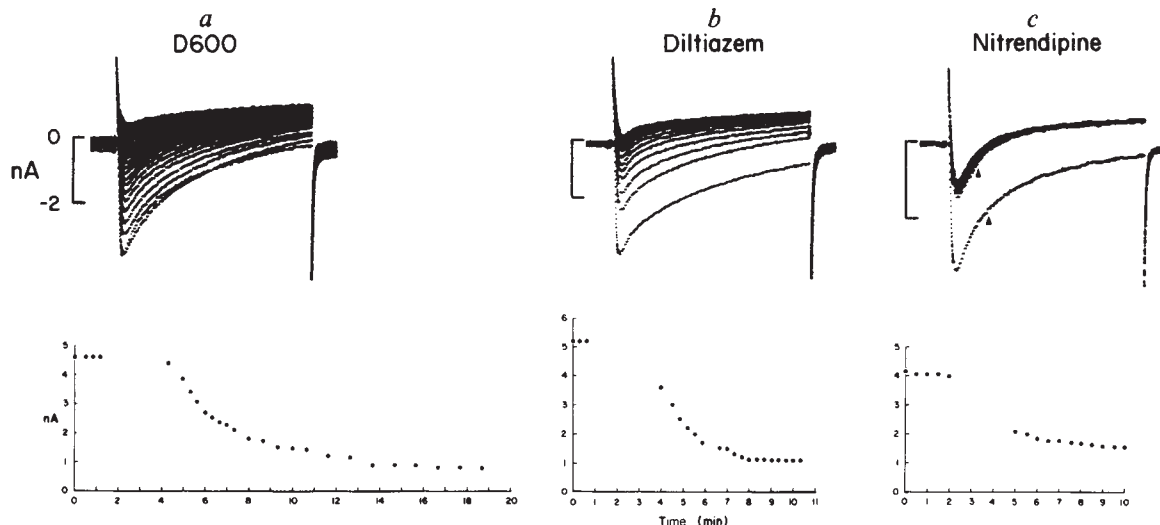


Fig. 5 Effect of repetitive depolarization on the development of blockade. *a*, $0.3 \mu\text{M}$ D600; *b*, $5 \mu\text{M}$ diltiazem; *c*, $1 \mu\text{M}$ nitrendipine. Shown above are records of total membrane current during 120-ms pulses to $+20$ mV from a holding potential of -40 mV. 10-Ca, 0-Na, 0-K solution outside, 151-Cs solution inside, as described in the text. In each panel, the largest inward current was recorded just before applying the drug; successively smaller Ca currents were recorded during a series of test pulses at $\sim 3 \text{ min}^{-1}$ starting 3 min after beginning the drug exposure. The graphs plot the magnitude of each of the peak Ca currents, measured after correction for linear capacity and leak currents. Cells 139A, 142D and 151B in *a-c*, respectively.

indicates that D600, diltiazem and nitrendipine fall along a spectrum. In Fig. 5a, 0.3 μM D600 produces very little inhibition of Ca current in the absence of test pulses, but blockade increases with repeated depolarizations over a period of many minutes. The cumulative effect of stimulation is an effect of the drug; in control runs (not shown here) the Ca current is unchanged during repetitive test depolarizations at 3 min^{-1} after a period of quiescence. Studied with the same pulse protocol, nitrendipine blockade is strikingly different (Fig. 5c): the very first current following the quiescent period gave a good estimate of the final level of blockade. This was a consistent finding for nitrendipine (see also Fig. 6a, b); at 3 min^{-1} , progressive changes in Ca current were small and could be explained simply by Ca channel run-down. The contrast between D600 and nitrendipine is in line with a previous comparison between blocking effects of D600 and nisoldipine by Scheuer and Kass²⁷. Figure 5b shows use-dependent effects of diltiazem, also consistent with earlier reports^{26,28,29}. We found that diltiazem was intermediate between D600 and nitrendipine in its use-dependence at drug doses giving similar steady levels of blockade.

Do the results with D600 and nitrendipine reflect a qualitative difference in mechanism? In previous comparisons, nifedipine was categorized as a drug that merely reduces I_{Ca} , while D600 was taken as an example of a drug that exerts additional effects on the kinetics of inactivation or repriming^{4,31}. We looked for effects of nitrendipine on Ca channel repriming, using pairs of depolarizing pulses separated by a variable interval. Nitrendipine promotes the appearance of two distinct phases in the time course of recovery: a relatively rapid recovery phase like that seen in the absence of drug, and a clear second phase with a time constant of several seconds at -40 mV. Thus, nitrendipine shows a qualitative resemblance to previously reported effects of D600^{25,27,29} in slowing Ca channel repriming. However, the time constant of the D600-induced slow phase is several minutes, not seconds, at comparable membrane potentials (refs 25, 29; Fig. 5a). Therefore, it is not surprising that pulses separated by 20 s produce substantial cumulative block with D600, but not with nitrendipine.

Blockade of open channels

Do drug molecules interact preferentially with activated (open) Ca channels? In the case of Ca channels, there is little or no direct evidence for this mechanism as opposed to blockade of inactivated channels. One clue to the possible importance of open channel blockade is illustrated in Fig. 5c: nitrendipine not only reduces the peak magnitude of the Ca current but also hastens its decay. The arrows marking the half-time of the relaxation ($t_{1/2}$) show a decrease in $t_{1/2}$ from 24 ms to 17 ms. The nitrendipine-induced acceleration is even more dramatic when Ca channel current is carried by Ba^{2+} rather than Ca^{2+} (Fig. 6a). During a 120-ms depolarizing pulse, Ba current shows very little inactivation in the absence of drug, but displays a prominent relaxation after exposure to nitrendipine. Some inactivation of the Ba current can be seen if the depolarizing pulse is prolonged to 600 ms (Fig. 6b); however, the speeding of the inward current relaxation by nitrendipine is still evident. Changes in the relaxation of inward current could also be seen with D600 during relatively long-lasting depolarizations in Ba (Fig. 6c). The shutting-off of the Ba^{2+} current is markedly accelerated during the first test pulse following the 3-min drug exposure. As blockade accumulates with subsequent depolarizations, the inward decay becomes less rapid. The initially rapid decay could be interpreted as drug blockade of open Ca channels. The gradual lessening of the acceleratory effect with repeated depolarization is also understandable because fewer unblocked channels are available for drug interaction as the level of blockade deepens towards its dynamic steady state.

We suggest that these effects may be direct expressions of the drug binding to activated Ca channels. Organic Ca channel blockers may also show preferential binding to inactivated Ca channels²⁵⁻²⁹, as in the case of local anaesthetics and Na

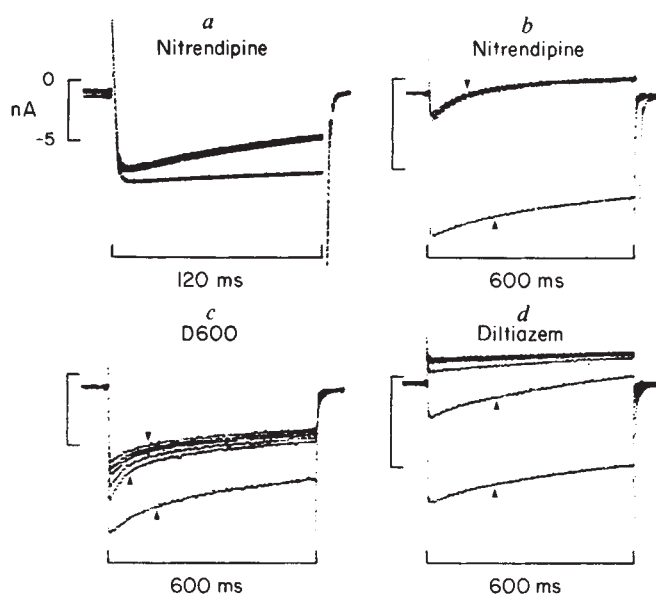


Fig. 6 Changes in the time course of Ba current through Ca channels, consistent with open channel block. Total membrane current recorded with 10-Ba, 0-Na, 0-K solution outside, 151-Cs solution inside. The experimental procedure was as described for Fig. 5a-c. In each panel, the largest inward current was total membrane current recorded with a test depolarization from -40 mV to $+20$ mV just before application of drug. Records of Ba current, taken during repeated depolarizations at 3 min^{-1} after 3 min drug exposure, were illustrated as follows: a, first six records after 1 μM nitrendipine; b, first, third, fifth and seventh records after 5 μM nitrendipine; c, first, third, fifth, seventh and ninth records after 0.3 μM D600; d, first six records after 50 μM diltiazem. Note that time scale is fivefold slower in b, c and d than in a. Vertical arrows mark half-completion of the decay of inward current that occurs during the pulse. The faster relaxation in a-c is consistent with drug-induced blockade of open channels but not with certain other kinetic interpretations such as drug binding to inactivated channels only. For example, suppose that inactivation takes place by a first order reaction with k_f and k_b as forward and backward rate constants, respectively, and $1/(k_f + k_b)$ as the observable time constant of relaxation. If drug bound exclusively to inactivated channels and prevented the back reaction ($k_b = 0$), the relaxation time constant would become $1/k_f$. Thus, the drug would slow relaxation, rather than speed it as actually observed.

channels^{32,33}. Further analysis is needed to determine the relative importance of these mechanisms for each of the inhibitors. Blockade of inactivated channels may be particularly important for diltiazem²⁶ because 'open channel block' is not prominent even at the relatively high concentration of 50 μM (Fig. 6d).

Concentration dependence of nitrendipine block

Measurements of ^3H -nitrendipine binding to membrane fragments from heart or other tissues demonstrate a saturable component of binding with an apparent K_D of <0.3 nM (see, for example, refs 6-9). However, sub-micromolar doses of nitrendipine are required to produce substantial blockade of Ca or Ba currents (Figs 2-4). Figure 7 shows pooled results from a large number of cells in 10 mM Ca_o at a holding potential of -40 mV. The data can be fitted fairly well by a one-to-one binding curve, with half-maximal block at ~ 150 nM. Our results seem consistent with the dose-dependent inhibition of contractile activity in papillary muscle by various nifedipine analogues³⁴. Thus, there is a 100-1,000-fold discrepancy in the drug concentrations required for physiological effects in functional cells and for radioligand binding to membrane fragments. One possibility is that high-affinity binding to broken membranes reflects the behaviour of Ca channels maintained in activated or inactivated states.

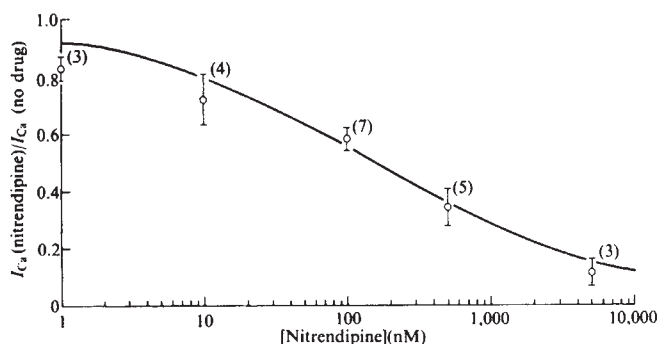


Fig. 7 Dose-dependence of Ca current block by nitrendipine. Ordinate: peak Ca current magnitude after 5–10 min drug exposure is plotted as a fraction of the peak Ca current before drug. Measurements after subtraction of leak and capacity currents. 10-Ca, 0-Na, 0-K solution outside, 151-Cs solution inside. Plotted are mean $\pm$ s.e.m. with the number of cells at each drug concentration given in parentheses. The smooth curve describes one-to-one binding with a dissociation constant of 154 nM.

Conclusions

Although diverse in structure, verapamil, D600, diltiazem and nitrendipine share several fundamental characteristics. They block outward currents through Ca channels, but leave the activation of unblocked channels essentially unchanged. The percentage of unblocked channels is increased by elevating

extracellular Ca^{2+} , in accordance with a genuine Ca antagonism. Thus, the classifying terms Ca channel blocker or Ca antagonist seem experimentally justified. One unexpected finding was that the antagonism between permeant ions and organic inhibitors cannot be explained by Ca^{2+} and Ba^{2+} acting at the same binding site as that at which inorganic ions block²³. This raises the possibility that the Ca channel contains more than one cation binding site, as also suggested by evidence for ion-ion interactions within the Ca channel³⁵.

The various organic compounds show interesting quantitative differences in the extent to which they display use-dependence. The effects of previous stimulation are most persistent in the case of verapamil, D600 and diltiazem, which are tertiary amines, and much less persistent for nitrendipine, which is uncharged at physiological pH. This pattern is reminiscent of differences between tertiary amine local anaesthetics (lidocaine, quinidine) and uncharged local anaesthetics (benzocaine) in the use-dependence of their blocking effects on Na channels^{32,33}. The potency of verapamil will be enhanced relative to that of other drugs by use-dependent blockade in rapidly firing tissue. Thus, differences in use-dependence may help explain why verapamil is much more cardioactive than nifedipine at doses equally effective on vascular smooth muscle.

We thank Esther W. Lee for technical assistance and Bruce P. Bean, Charles J. Cohen and Ronald A. Janis for helpful discussion. This work was supported by grants from the Miles Institute, Marion Laboratories, the American Heart Association and the USPHS.

Received 8 November 1982; accepted 15 February 1983.

1. Fleckenstein, A. A. *Rev. Pharmac. Tox.* **17**, 149–166 (1977).
2. Katz, A. M. & Reuter, H. *Am. J. Cardiol.* **44**, 188–190 (1979).
3. Trigg, D. J. in *Calcium Blockers: Mechanisms of Action and Clinical Applications* (eds Flaim, S. F. & Zelis, R.) 121–134 (Urban & Schwarzenberg, Baltimore, 1982).
4. Henry, P. D. *Am. J. Cardiol.* **46**, 1047–1058 (1980).
5. Antman, E. M., Stone, P. H., Muller, J. E. & Braunwald, E. *Ann. intern. Med.* **93**, 875–885 (1980).
6. Bellemann, P., Ferry, D., Lubbecke, F. & Glossmann, H. *Arzneimittel-Forsch.* **31**, 2064–2067 (1981).
7. Ehler, F. J., Itoga, E., Roeske, W. R. & Yamamura, H. *Biochem. biophys. Res. Commun.* **104**, 937–943 (1982).
8. Gould, R. J., Murphy, K. M. M. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3656–3660 (1982).
9. Bolger, G. T. *et al. Biochem. biophys. Res. Commun.* **104**, 1604–160 (1982).
10. Lee, K. S., Weeks, T. A., Kao, R. L., Akaiki, N. & Brown, A. M. *Nature* **278**, 268–271 (1979).
11. Brown, A. M., Lee, K. S. & Powell, T. J. *Physiol., Lond.* **318**, 455–477 (1981).
12. Lee, K. S. & Tsien, R. W. *Nature* **297**, 498–501 (1982).
13. Reuter, H. & Scholz, H. J. *Physiol., Lond.* **264**, 17–47 (1977).
14. Marban, E. & Tsien, R. W. J. *Physiol., Lond.* **329**, 569–587 (1982).
15. Nawroth, H., Ten Eick, R. E., McDonald, T. F. & Trautwein, W. *Circulation Res.* **40**, 408–414 (1977).
16. Noma, A. *et al. Nature* **285**, 278–279 (1980).

17. Isenberg, G. & Klöckner, U. *Nature* **284**, 358–360 (1980).
18. Powell, T., Terrar, D. A. & Twist, V. W. J. *Physiol., Lond.* **319**, 82–83P (1981).
19. Hescheler, J., Pelzer, D., Trube, G. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **393**, 287–291 (1982).
20. Kohlhardt, M., Bauer, B., Krause, H. & Fleckenstein, A. *Pflügers Arch. ges. Physiol.* **335**, 309–322 (1972).
21. Payet, M. D., Schanne, O. F. & Ruiz-Ceretti, E. J. *molec. cell. Cardiol.* **12**, 635–638 (1980).
22. Akaiki, N., Brown, A. M., Nishi, K. & Tsuda, Y. *Br. J. Pharmac.* **74**, 87–95 (1981).
23. Hagiwara, S. & Bylerly, L. A. *Rev. Neurosci.* **4**, 69–125 (1981).
24. Ehara, T. & Kaufmann, R. J. *Pharmac. exp. Ther.* **207**, 49–55 (1978).
25. McDonald, T. F., Pelzer, D. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **385**, 175–179 (1980).
26. Kanaya, S. & Katzung, B. G. *Circulation* **64**, IV-274 (1981).
27. Scheuer, T. & Kass, R. S. *Biophys. J.* **37**, 341a (1982).
28. Tung, L. & Morad, M. in *Ca²⁺ Entry Blockers, Adenosine, and Neurohumors* (ed. Merrill, G.) Ch. 2 (Urban & Schwarzenberg, Baltimore, 1983).
29. Pelzer, D., Trautwein, W. & McDonald, T. F. *Pflügers Arch. ges. Physiol.* **394**, 97–105 (1982).
30. Heistracher, P. *Naunyn-Schmiedeberg's Archs Pharmac.* **269**, 199–212 (1971).
31. Kohlhardt, M. & Fleckenstein, A. *Naunyn-Schmiedeberg's Archs Pharmac.* **298**, 267–272 (1977).
32. Hille, B. J. *gen. Physiol.* **69**, 497–515 (1977).
33. Hondeghem, L. M. & Katzung, B. G. *Biochim. biophys. Acta* **472**, 373–398 (1977).
34. Rodenkirchen, R. *et al. Naunyn-Schmiedeberg's Archs Pharmac.* **310**, 69–78 (1979).
35. Hess, P., Lee, K. S. & Tsien, R. W. *Biophys. J.* **41**, 293a (1983).

Complete amino acid sequence of a mouse epidermal keratin subunit and implications for the structure of intermediate filaments

Peter M. Steinert*, Robert H. Rice*, Dennis R. Roop†, Benes L. Trus‡
& Alasdair C. Steven§

* Dermatology Branch, and † Laboratory of Cellular Carcinogenesis and Tumor Promotion, National Cancer Institute, NIH, Building 10, Rm 12N238, Bethesda, Maryland 20205, USA

‡ Computer Systems Laboratory, Division of Computer Research and Technology, and § Laboratory of Physical Biology, National Institute of Arthritis, Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda, Maryland 20205, USA

We have determined the complete primary structure of an intermediate filament subunit, the 59,000 molecular weight subunit of mouse epidermal keratin, from the nucleotide sequence of cDNA clones. The central portion of the sequence forms extended tracts of a coiled-coil α -helical conformation. This is flanked at both termini by similar non- α -helical sequences that are extremely rich in glycine residues, frequently configured in tandem peptide repeats. Limited chymotryptic digestion of keratin filaments containing this protein suggests a structural organization whereby the terminal glycine-rich sequences protrude from a conserved core structure into which the coiled-coil α -helical segments are packed.

INTERMEDIATE filaments (IF) are ubiquitous constituents of the cytoskeleton of eukaryotic cells. They consist of a heterogeneous population of subunit types, which on the basis of their tissue or cell type of origin, variations in immunological and solubility properties, and conditions for self-assembly *in vitro*, have been operationally categorized into five distinct subclasses: the keratins or cytokeratins of epithelia; neurofilaments of neuronal tissues; muscle desmin; astrocyte glial fibrillary acidic protein; and vimentin of mesenchymal cells^{1,2}. While the subunits of IF vary widely in these properties, they share certain common structural features: all form IF *in vivo* and *in vitro* that are morphologically similar³, all seem to contain similar peptide domains^{4,5}, and all contain extensive α -helical sequences that assume a coiled-coil conformation; that is, they are α -type fibrous proteins of the k-m-e-f class⁶. Although little is known about the principles of IF structure which allow the observed variations but nevertheless maintain the perceived similarities, it is apparent that fundamental structural information evinced for one type of IF should be of direct relevance to all other types of IF.

We report here the complete amino acid sequence of an IF subunit of mouse epidermal keratin. Newborn mouse epidermis synthesizes several keratin IF subunits of two major size classes: a high molecular weight (M_r) 'K₁' family consisting of proteins of 68,000 (68K) and 67,000 (67K) M_r ; and a lower molecular weight 'K₂' family consisting of proteins of 60K and 59K M_r (refs 7-9). Differences in the temporal expression of the members of these two families in cells of intact epidermis and cells grown in culture have been reported^{7,9}. Members of both families are required for keratin IF assembly *in vitro*⁸. Recently, a library of cDNA clones was prepared from poly(A) RNA of newborn mouse epidermis⁹. Two of these recombinant plasmids, pK 276 (0.536 kilobase pairs, kbp) and pK 435 (1.802 kbp) were shown by hybrid selection analysis to contain cDNA sequences complementary to the mRNA encoding the 59K keratin⁹. We describe here the nucleotide sequence of these clones and the complete deduced amino acid sequence of the protein. The structural implications of this sequence information to IF structure are developed below.

Nucleotide sequence and assignment of protein sequence

The complete DNA sequences of the pK 276 and pK 435 clones were determined (Fig. 1). Primer extension analysis indicated that the larger clone, pK 435, lacked ~145 nucleotides from the 5' end of the mRNA for the 59K protein. To obtain the remaining 5' sequence, a 5'-end-labelled *Dde*I to *Sac*I primer was isolated, hybridized to newborn mouse epidermal poly(A) RNA and extended by reverse transcriptase as previously described¹⁰. The extended primer was then sequenced directly. The completed nucleotide sequence revealed a unique initiating ATG codon near the 5' end but did not possess a poly(A) tail or putative poly(A) addition sequence near the 3' end. These sequences were presumably lost during the preparation of the cDNA clones. The nucleotide sequence was characterized by two long stretches of nucleotides (roughly, nucleotides 50-400 and 1,350-1,680) that were extraordinarily rich in the base G. These regions were difficult to sequence by the Maxam-Gilbert¹¹ and other procedures involving denaturing gel electrophoresis on account of the paucity of suitable restriction enzyme sites and also because of the refractory secondary structure resulting from these sequences.

Unique assignment of an amino acid sequence was readily possible because the codon frame containing the ATG initiating triplet corresponded to the G-rich frames that encode the large numbers of glycine residues known to be present in the protein⁸ and to the only sense coding frame throughout the entire nucleotide sequence. The deduced sequence contained a total of 569 amino acids, giving a calculated molecular weight of 57,850. As this protein contains about 5 mol per mol of *O*-phosphoserine¹², the true M_r of the protein is 58,350, which is

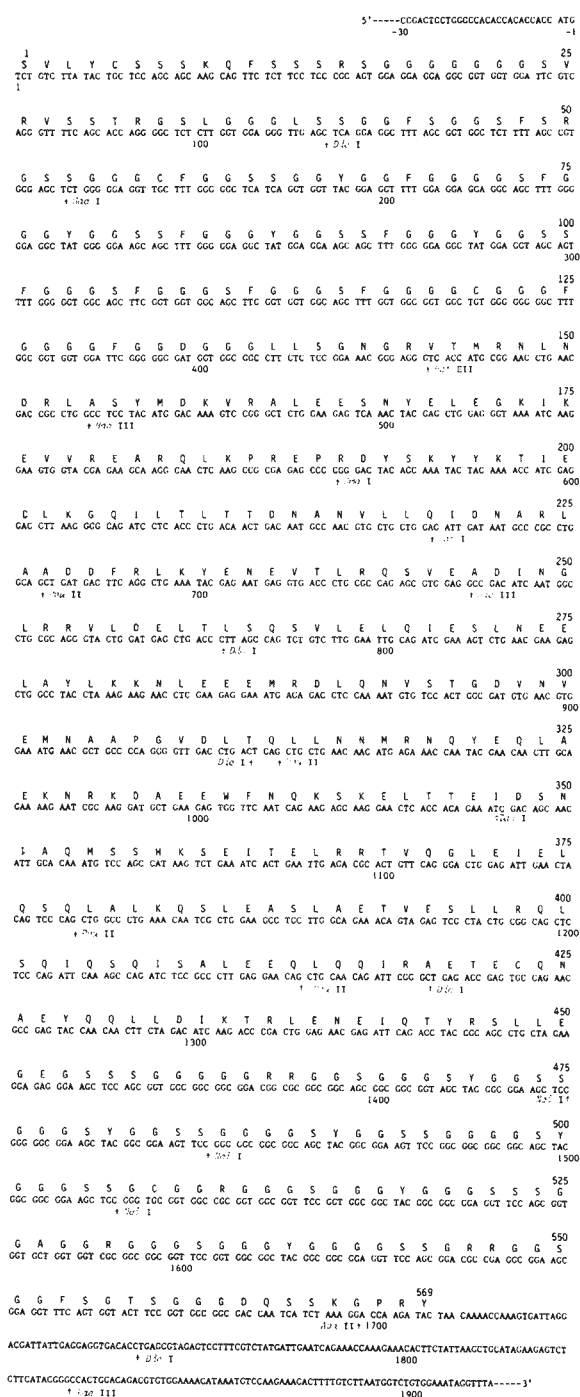
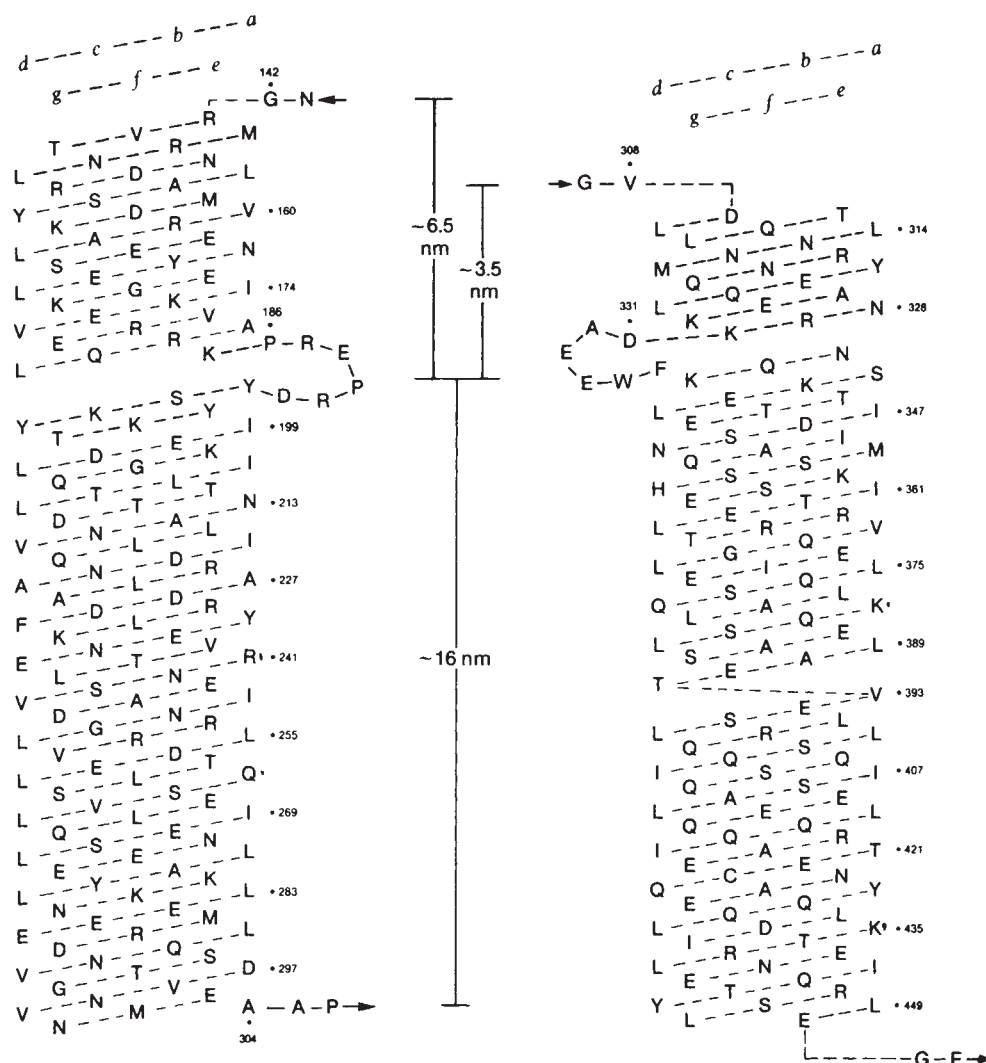


Fig. 1 Primary structure of the mouse epidermal keratin subunit. The nucleotide sequence was determined using a restriction endonuclease map constructed in preliminary experiments. The figure displays only those sites actually used in sequencing analysis. Sequencing was carried out following 5' and 3' labelling of appropriate fragments on both strands using the Maxam and Gilbert¹¹ procedures. The clone pK 276 (0.536 kbp) began just upstream from the first *Nci*I site and pK 435 (1.802 kbp) began two bases upstream from a *Dde*I site. The 5'-terminal sequence was then derived following primer extension¹⁰ of a probe consisting of a *Dde*I-*Sac*I fragment. Thus, the 5' sequence may not extend to the true 5' end of the mRNA. Nucleotides are numbered in the 5' to 3' direction, beginning with the first base following the initiating triplet ATG. The poly(dC)-poly(dG) tracts derived in cloning are not included. The 3'-terminal sequence shown does not extend to the poly(A) tract (see text). The deduced amino acid sequence is displayed above the nucleotide sequence using the single-letter amino acid code. The amino acids are numbered from the serine residue following the initiating codon, as a previous study⁸ has shown that *N*-acetylserine is the amino-terminal residue of K₂ proteins.

Fig. 2 Amino acid sequence of the α -helical portions of the subunit arranged as a helical coiled-coil¹⁷. The sequences are displayed as though situated on a cylinder which has been cut open and then spread out flat. The α -helix has 3.6 residues per turn in the coiled-coil, but for clarity, is drawn here as alternating offset rows of 4 and 3 residues so that the α -helix axis seems to rotate at the same rate as the axis of the coiled-coil. The coiled-coil on the left is constructed from the first portion of the α -helical sequences. It is interrupted by a non- α -helical inclusion of about six residues that would be folded out of the coiled-coil due to its content of two prolines. The length of this coiled-coil is about 23 nm, in two segments of 6.5 and 16 nm. The coiled-coil on the right is from the latter portion of the α -helical sequences and has been separated from the left by residues 306–307 that form a part of a second non- α -helical interruption. This coiled-coil region is interrupted by a third discontinuity at residues 330–336. Later, the coiled-coil is disrupted between Thr 392 and Val 393 by a reversal in the polarity of the a and d heptad designation, which may result in a rotation of the backbone of the coiled-coil (see text). The length of this coiled-coil portion is about 20 nm, in two segments of 3.5 and 16 nm, giving a total length of about 42 nm for the entire coiled-coil α -helical region of the keratin subunit. The determination of the exact amino acids which enter and leave the coiled-coils, especially at doubtful positions such as Asp 297, will require more comprehensive model-building studies.



very close to the mass of 59,000 estimated by SDS-gel electrophoresis. Other criteria which confirm this assignment of amino acid sequence are: the deduced amino acid composition is very similar to that previously reported⁸ for K₂ proteins (for example, 25.8% glycine content compared with a reported value of 25.9%); several peptides generated by limited chymotryptic digestion of keratin IF containing this protein had sequences identifiable in the protein (see below); the structural features evident from this sequence are also consistent with known properties of the protein (see below).

Structural features of the amino acid sequence

The availability of the amino acid sequence permits use of computer-assisted analyses by the Chou and Fasman¹³ and Robson¹⁴ methods, to identify regions that potentially form α -helices, β -sheet structures and β -turns, which often demarcate the structural domains of a protein. All of the α -helix of this mouse keratin protein is located in a central region, bounded approximately by residues 143 and 450, with several interruptions by prominent β -turns near residues 185–191, 305–309 and 315–320. These methods of analysis indicate that the amino- and carboxy-terminal sequences, which are rich in glycine residues, are folded about 100 times each by β -turns. However, because such analytical methods cannot readily be applied to fibrous proteins, we have undertaken more detailed

examination of the amino acid sequence by the Fourier transform method¹⁵ to identify the nonrandom distributions of certain types of amino acids, and have found highly significant periodicities in a variety of amino acids (Table 1). From these periodicities, precise structural predictions can be made about the α -helical and glycine-rich sequences, as described below.

α -Helical sequences

The Fourier transform periodicities of 7/2 and 7/3 in apolar amino acids (Table 1) are clearly harmonics of a 7-residue repeat characteristic of α -helical coiled-coils¹⁶, such as have previously been identified in α -type proteins like tropomyosin^{17,18}, myosin rod¹⁹, fibrinogen²⁰ and wool keratin^{21,22}. The α -helical sequences can be displayed two-dimensionally as a coiled-coil net¹⁷ consistent with the helical arrangement of two or three adjacent strands of α -helix. To form a stable coiled-coil, residues in the first (a) and fourth (d) positions of the heptad should have apolar side chains that form a closely packed hydrophobic interface between the strands of the coiled-coil¹⁶. Residues b , c , e , f and g occupy superficial positions in the coiled-coil and are usually polar or charged¹⁶. In the coiled-coils we have constructed for this mouse epidermal keratin subunit (Fig. 2), 31 of 40 residues in the a positions and 33 of 39 residues in the d positions are occupied by apolar amino acids; this is highly consistent with the above model. Amino acids with long side chains, such as lysine, occur seven times in these

Table 1 Significant repeats in the sequence from Fourier analysis

Residue type/ subsequence	Period	Rational approximation	Probability of significance (>0.95)
Apolar (AVLIMFY)			
192-293	3.50	7/2	0.963
	2.33	7/3	0.958
336-450	3.52	7/2	0.978
Acidic (DE)			
192-293	9.33	28/3	0.971
336-450	2.81	28/10	0.951
Basic (KR)			
192-293	9.46	28/3	0.968
	7.07	7 or 28/4	0.956
336-45	8.64	26/3	0.955
F and Y			
1-142	4.79	19/4	0.981
451-569	3.17	19/6	0.950
Glycine (G)*			
1-142	4.76	19/4	0.974

Preliminary Fourier analyses of the total sequence gave strong indications of these repeats. However, the more refined analyses tabulated here associate the repeats with specific regions of the molecule. For short sequences, fluctuations can occur on account of coarse sampling of the Fourier transform: consequently, all sequences analysed were padded with zeros to give a total of 700 samples. The analysis was performed and probabilities inferred from the intensities of the Fourier components according to McLachlan¹⁵.

* The C-terminal sequence (451-569), which is closely homologous to the N-terminal sequence, also indicated a similar G repeat of 19/4 residues, although with a reduced probability of 0.93.

positions but are unlikely to distort the coiled-coils significantly¹⁶; acidic and short-side-chain amino acids that would introduce distortions¹⁶ are rare. Of the remaining 192 superficial positions on the coiled-coils, 148 are occupied by hydrophilic amino acids, of which 75 are charged (Fig. 2).

The coiled-coils shown in Fig. 2 represent a plausible deployment of the expected secondary structure, but it should be noted that the precise arrangements will require higher resolution structural studies that take into account the interactions permitting the formation by these sequences of either two-chained²³ or three-chained^{4,24} coiled-coils. The first region consistent with a coiled-coil begins at Arg 143 and extends to Lys 185 where it is interrupted by a sequence containing two prolines which form β -turns and thus a globular loop folded out of the coiled-coil (Fig. 2). The coiled-coil resumes at Tyr 192 and extends for about 15 heptads to Val 293. At least part of the sequence Ser 294 to Ala 304 could also form a coiled-coil but Asp 297 in the *a* position would impose a significant distortion¹⁶. The coiled-coil is clearly terminated by another globular loop around Pro 306. The conformation assumed by residues 309-335 is equivocal. It is possible that the coiled-coil resumes at Asp 309 and extends for 3 heptads to Lys 330 where it is again interrupted by irregularities between Asp 331 and Tr 335. Alternatively, the short length of this piece and its proximity to irregular sequences may render it unstable so that instead of a coiled-coil this region adopts a globular but α -helix-rich conformation. Another stable coiled-coil forms between Glu 336 and 450. However, the polarity of the *a* and *d* heptad designation of this coiled-coil reverses between Thr 392 and Val 393, a phenomenon already encountered in the partial sequences of the IF proteins desmin and vimentin⁵. This discontinuity probably results in a rotation of the backbone of the coiled-coil at this point.

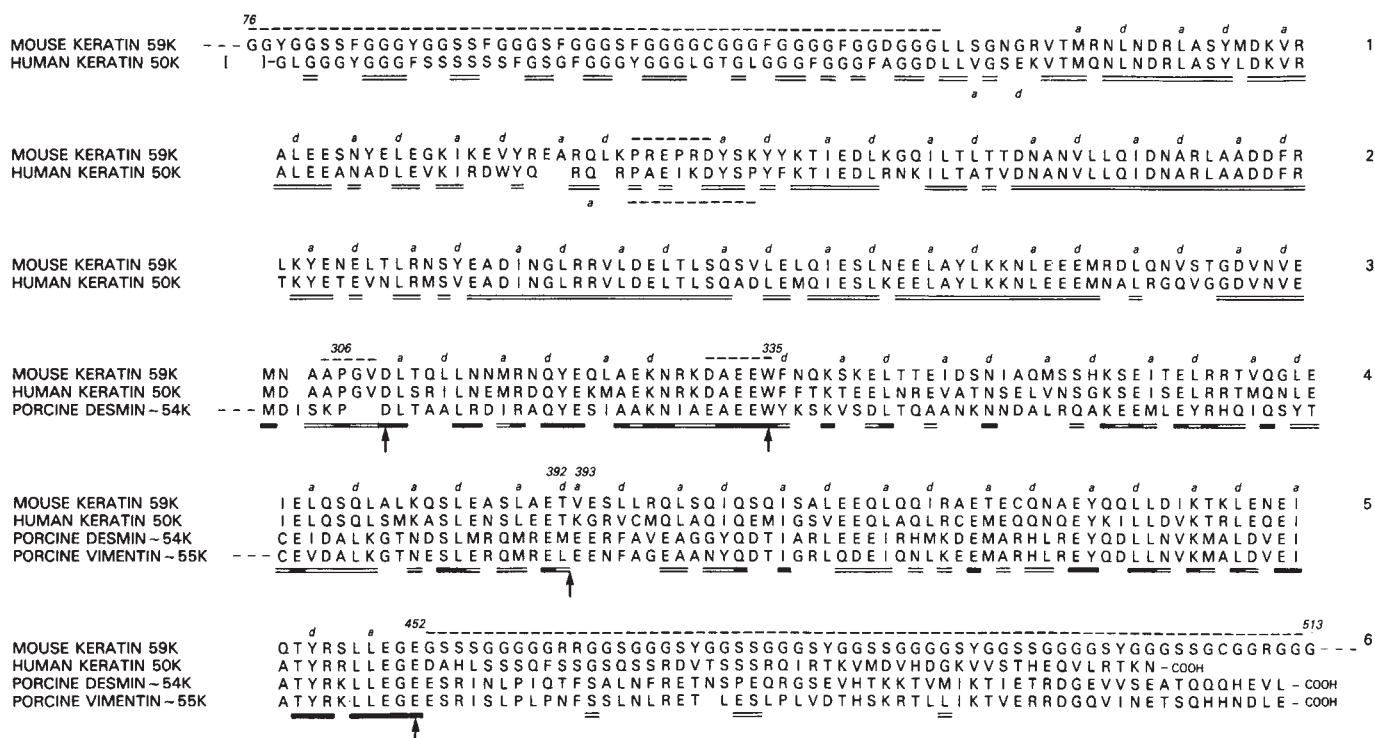


Fig. 3 Sequence homologies of keratin, desmin and vimentin IF subunits. The mouse keratin sequence shown extends from Gly 76 to Gly 511 and is aligned to maximize homologies with the available sequence information of a human keratin subunit²⁶ and porcine desmin and vimentin^{5,27,28} which extend to their carboxy-termini. The *a* and *d* heptad designations are given. Arrows mark key positions of identity: proline (mouse keratin residue 306), tryptophan (residue 335), the point of heptad polarity reversal (residues 392,393), and the end of the helical coiled-coil (near Glu 452). Underscored amino acids indicate positions of identity between the human and mouse keratins; amino acids underscored in dark indicate positions of identity in the four proteins. Many other positions show concurrence of homologous amino acids, such as interchanges of various apolar residues, for example lysines for arginines, which might not be expected to alter the structure of the coiled-coils. Regions which assume a non- α -helical conformation are denoted by broken lines. Comparisons with other published IF subunit sequences are not shown because of the limited data available.

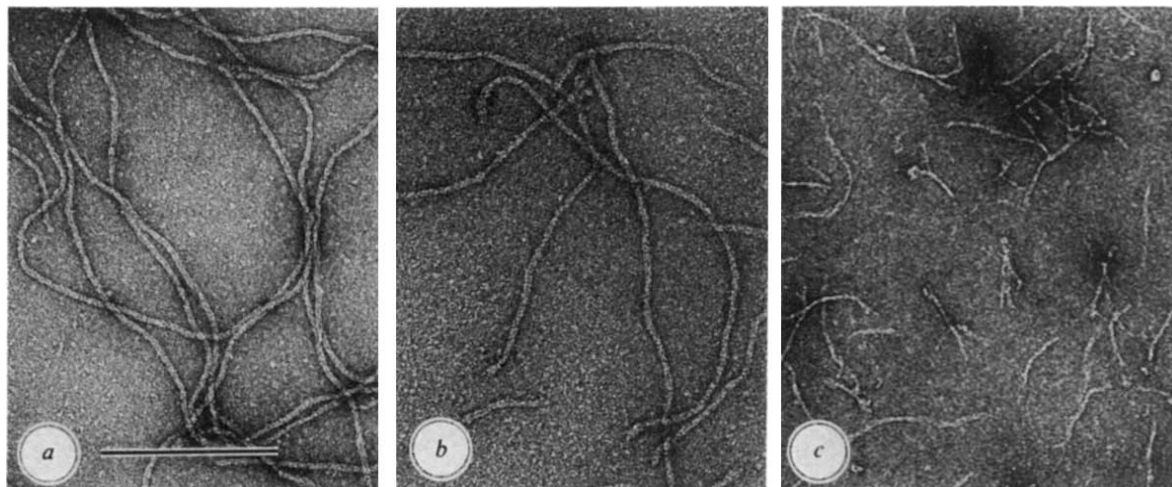


Fig. 4 Structure of keratin IF following chymotrypsin digestion. Freshly isolated newborn mouse epidermal cells, which actively synthesize keratin subunits⁸, were cultured (10^7 cells ml^{-1}) for 1 h in medium containing $10 \mu\text{Ci ml}^{-1}$ each of ^3H -glycine and ^{32}P -orthophosphate (NEN). The cells were collected, washed and extracted with urea buffer to isolate the keratins, which were purified by column chromatography on DEAE-cellulose⁸ and polymerized in assembly buffer at $550 \mu\text{g ml}^{-1}$ into keratin IF *in vitro*^{3,8}. About 96% of the total protein in solution could be pelleted at $100,000g$ in 30 min (ref. 8), indicating a high yield of protein in IF. As determined by negative staining, the filaments were more than $2 \mu\text{m}$ long and uniformly $8\text{--}10 \text{ nm}$ wide (a). Samples of these IF were then exposed to chymotrypsin ($0.5 \mu\text{g ml}^{-1}$, 0.1% w/w) for varying times. By 10 min, the IF still has a 'normal' appearance as determined by negative staining, although they were somewhat shorter ($\sim 1 \mu\text{m}$ long) (b). However, only about 68% of the original protein in solution could be pelleted, indicating that about one-third of the mass had been removed from the IF. By 30 min of digestion, the IF had been reduced to α -helix-enriched particles that could no longer be pelleted (c), as we have described previously²⁴.

Thus, a virtually continuous coiled-coil can be constructed from Arg 143 to Glu 450, with three short interruptions. Because each amino acid in an α -helix assumes an axial displacement of about 0.15 nm , the entire α -helical portion of this mouse epidermal keratin subunit is about 42 nm long, in discrete segments of 6.5 , 16 , 3.5 and 16 nm . The contiguous length assumed by the globular loops is unknown, but their small size and probable structure suggest that they contribute little if anything to the overall length of the coiled-coil α -helical region; perhaps a maximum of another 3 nm may be possible (assuming 0.2 nm per residue). This length of $42\text{--}45 \text{ nm}$ is quite consistent with a previous estimate of 40 nm (refs 4, 24). However, the structural arrangement of the coiled-coils is now evidently different from our earlier predictions. The limited proteolytic digestion procedures used earlier presumably cleaved the intervening non- α -helical globular loops, yielding only the two longest coiled-coils; the smaller coiled-coils were apparently lost during digestion.

The Fourier transform analyses (Table 1) also revealed highly significant periodicities of $28/3$ and $28/10$ in the distributions of acidic and basic residues in the α -helical regions. A repeat of 28 residues corresponds to an axial extension of about 4.2 nm of the coiled-coils. Such a repeat is also evident in the available sequences of vimentin and desmin⁵. As the charged residues occupy superficial sites of the coiled-coils, they probably have a very significant role in the lateral association of neighbouring coiled-coils in the IF structure. Indeed, our more extensive analyses of the sequence data (P.M.S., in preparation) suggests that these charged interactions favour the parallel packing of neighbouring coiled-coils in approximate alignment with the filament axis. In this regard, it is noteworthy that the $28/3$ and $28/10$ repeats are absent from the sequences of tropomyosin, the coiled-coils of which do not normally form lateral arrays.

Glycine-rich sequences

The 10–15 amino acids located on the amino- and carboxy-terminal ends of the subunit contain a variety of amino acids that favour the assumption of β -turns. Immediately adjacent to these residues, there occurs a strikingly high incidence of consecutive glycine residues, and later, tandem peptide repeats. Examples of such repeats are: FG GYG GSS (residues 74–82, 83–91, 92–100), followed by FG GGS (residues 101–105,

106–110, 111–114), and $\text{G}_3\text{C}_2\text{GGRGGGSGGGYGGGGSSG}$ (residues 506–525, 526–545). Although the exact nature of the repeating sequences at the respective ends of the subunit vary, Fourier analyses suggest that the two glycine-rich sequences are evolutionary related ($<10^{-23}$ chance of having arisen independently) and have similar structures and functions within the keratin IF. Both glycine-rich sequences abut directly onto the coiled-coil α -helical sequences.

The Fourier analyses (Table 1) illustrated highly significant periodicities of about $19/4$ and $19/6$ in the distributions of glycines and phenylalanines + tyrosines in the two glycine-rich regions, and visual appraisal of the sequences indicates that tyrosines and phenylalanines repeat at every fourth or fifth residue, with an average of 4.75 residues. However, the relevance of a 19-residue repeat to keratin IF structure is not yet clear. The sequences contain no α -helix, but short peptides such as F_3GGSS are suggestive of the formation of a β -conformation, whereas sequences involving longer runs of consecutive glycines are likely to introduce frequent folds in the protein backbone. Several short segments of a polyglycine type II structure may be possible between Gly 75 to Gly 135 and Gly 457 to Gly 528. Such a structure requires a glycine in every third position and forms a three-chain rod structure²⁵. We postulate that these glycine-rich sequences are highly convoluted and flexible. Keratins in general contain high contents of glycine residues^{3,8} and our nucleotide sequence analyses of other keratin clones indicate that glycine residues are also concentrated in the non- α -helical terminal regions (P.M.S. *et al.*, unpublished data). Such sequences may be characteristic of the epidermal keratin (or cytokeratin) type of IF and represent a hitherto undescribed structural form.

The glycine-rich sequences must be required for some aspect of keratin IF assembly or elongation of protofilaments to form the IF, as previous protein chemical studies have shown that the coiled-coil α -helical segments isolated on proteolytic digestion of intact keratin IF (or their protofilaments) had no potential for assembly into filamentous structures *in vitro* (ref. 24 and P.M.S., unpublished data).

Comparisons with other IF subunits

The latter half of the α -helical sequence of the mouse keratin subunit shows substantial homologies with the published

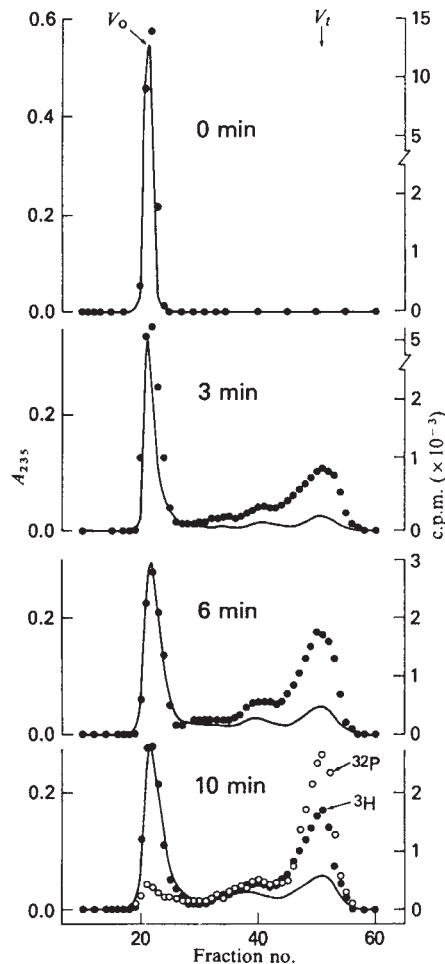


Fig. 5 Column chromatography of chymotrypsin digestion products. The protein was obtained from digestions as described in Fig. 4 legend. At the times indicated, 0.5-ml aliquots were removed, mixed with 0.1 vol. of 2% SDS, 2% dithiothreitol and 1% phenylmethylsulphonyl fluoride, heated for 2 min at 90 °C and then applied to a 0.9 × 95 cm column of Biogel P10 (Biorad) equilibrated in 50 mM Tris-HCl (pH 8.0), 2 mM dithiothreitol, 0.1% SDS. Aliquots were removed for determination of the distribution of radioactivity. Some of the excluded protein from each column run was used for SDS-gel electrophoresis and the remainder was used for determination of α -helix content by circular dichroism³. The α -helix content of the protein material at 0, 3, 6 and 10 min of digestion was: 39 ± 5%, 47 ± 4%, 57 ± 5% and 62 ± 3%, respectively. In a similar experiment, glycine-rich peptides obtained at 10 min of digestion of 5.7 mg of keratin IF were collected and freed of SDS. By measurement of radioactivity, 70% of the ³H- and 98% of the ³²P-labelled peptides were insoluble in aqueous solvents, but 30% of the ³H label was soluble in pyridine/acetic acid/water (1:10:289, v/v/v, pH 3.5) buffer. These basic peptides were separated on a thin layer sheet by electrophoresis (1,000 V, 30 min) in this solvent and then resolved in the second dimension by ascending chromatography using 2.5% (w/v) formic acid. The dried thin layer sheet was treated with Enhance (NEN) and exposed to X-ray film. Ten well separated labelled peptides were eluted and dried. Some of each eluate was subjected to acid hydrolysis to determine the amino acid compositions. The remainder was digested with carboxypeptidases A and B (Sigma) or carboxypeptidase Y (Calbiochem) and the products were assayed by amino acid analysis⁴. Complete or near complete sequencing was thus possible. Three peptides had sequences identifiable with the 59K sequence. Several other peptides have been localized to the 67K keratin subunit (P.M.S., in preparation).

sequences of the carboxy-terminal α -helical portions of a human epidermal keratin IF subunit of M_r 50,000 (ref. 26) and of porcine desmin and vimentin^{5,27,28} (Fig. 3, lines 4–6). Four major features of the sequences are located in identical positions: the interrupting non- α -helical globular loop, a unique

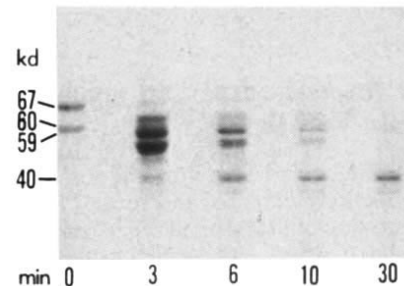


Fig. 6 Gel electrophoresis of chymotrypsin digestion products. Excluded protein from Fig. 5 was used. The original keratin IF used contained three subunits in approximate amounts by weight of: 67K, 50%; 59K and 60K, about 25% each.

tryptophan residue, the beginnings and ends of the coiled-coil regions, and even the position of reversal of heptad polarity. Whereas the mouse and human keratin subunits display 69% sequence identity and 79% sequence homology, and desmin and vimentin display 88% identity and 93% homology in these α -helical sequences, the keratins show about 34% identity and 51–54% homology to desmin. These features of relatedness indicate that this portion of the α -helix of these types of IF subunits are closely related structurally ($<10^{-17}$ chance of having arisen independently) and presumably functionally, although it is clear that the keratins diverged earlier in eukaryotic cell evolution from desmin and vimentin. Interestingly, however, such close homologies do not extend to larger keratin subunits. Partial sequence analyses of clones for the 67K mouse keratin show little sequence homology with the other IF subunits in the corresponding portion of the subunit (P.M.S. *et al.*, unpublished data).

The first half of the α -helical sequences of the mouse and human keratins are also very similar, but a significant variation occurs in the vicinity of the two proline residues (Fig. 3, line 2). The human keratin has three amino acid deletions and its two proline residues are more separated, thus forming a larger non- α -helical globular loop than in the mouse protein at this point. However, the coiled-coil of the human protein starts about one heptad earlier (Fig. 3, line 1) so that the overall length of the coiled-coils of the two proteins is about the same.

In striking contrast to the sequence similarities and homologies of the α -helical regions, the carboxy-terminal non- α -helical sequences of the different IF subunits are very different: the mouse keratin is uniquely glycine-rich (Fig. 3, line 6). These sequences may, however, adopt similar conformations, as the vimentin and desmin are likely to be folded many times at their proline residues and the human keratin at its serine residues by β -turns. The partial sequence information of the amino-terminal end of the human keratin shows that it is also glycine-rich (Fig. 3, line 1). Based on the known size of the full-length mRNA encoding this protein²⁶, we estimate that it contains an additional 40–50 residues, and based on the known amino acid composition²⁶, many of these will be glycines. Preliminary analysis of it reveals some sequence repeats and a 4.5–5-residue repeat in tyrosines and phenylalanines. Therefore, the amino-terminal regions of the mouse and human keratins may be structurally homologous. Overall, the human keratin is shorter by about 100 terminal non- α -helical residues than the mouse keratin.

In support of our earlier views^{2,4}, the available amino acid sequence data seem to suggest that the coiled-coil α -helical portions of different IF subunits are structurally homologous and of the same size, whereas the terminal non- α -helical sequences of the subunits are highly variable in size and amino acid sequence. More sequencing will be required to confirm this general scheme. Furthermore, as the human 50K keratin is a product of less differentiated epidermal cells and simple epithelia²⁹, and the mouse 59K keratin is expressed only in intact keratinizing epidermis⁹, such studies may also identify systematic sequence variations that correlate with the degree

of differentiation of the cells from which the IF subunits are obtained.

Location of the α -helical and glycine-rich regions within keratin IF

Several studies have suggested a non-uniform distribution of mass in IF^{24,30}. We have explored this further here by limited chymotryptic digestion of mouse epidermal keratin IF containing the 59K subunit which had been labelled *in vivo* with ³H-glycine and ³²P-orthophosphate. Whereas exposure of keratin IF to 1–2% (w/w) trypsin degrades the IF to α -helix-enriched particles within 2–10 min (ref. 24), exposure to 0.1% (w/w) chymotrypsin has little effect on the IF in this time as visualized by electron microscopy following negative staining (Fig. 4). However, analysis of the digestion products by other means demonstrated significant changes. Determination of mass in IF by ultracentrifugation analysis revealed that only 68% of the protein could be pelleted after 10 min of digestion, as opposed to 95% before digestion. Column chromatography in a denaturant solvent revealed that within 10 min about 68% of the total glycine content and more than 95% of the phosphate content of the IF had been removed as small peptides (Fig. 5). Gel electrophoresis of the excluded protein showed that all of the constituent subunits had been decreased in apparent molecular weight by 15–20,000 (Fig. 6). Estimation of the α -helix content of this excluded protein by circular dichroism revealed an increase from 39 $\pm$ 5% before digestion to 61 $\pm$ 4% after 10 min of digestion. Once separated, 70% of the glycine- and phosphate-labelled peptides included by the column were

insoluble in aqueous solvents, but 30% were cationic in charge and were soluble in pH 3.5 pyridine-acetate buffer. These were separated by two-dimensional peptide mapping and selected peptides were subjected to enzymatic sequencing analysis. Three peptides had sequences of GGGSSGRRGGSGGF, SGTSGGGDQSSKGPY and SRGSSGGGCF, which correspond exactly to residues 539–553, 554–569 and 49–58, respectively, in the 59K subunit.

These results are all consistent with the view that sequences containing at least two-thirds of the glycine and virtually all of the phosphate (as *O*-phosphoserine) of the subunits could be removed by very limited proteolysis, leaving an α -helix-enriched residue of morphological structure substantially identical to that of the original IF (Fig. 4). Thus, the glycine-rich sequences of the keratin IF subunits occupy more exposed or peripheral positions on the IF. The α -helical regions form the basic backbone or 'core' of the IF.

As illustrated in Fig. 3, the size, amino acid sequence and structure of the α -helical sequences of keratin, desmin and vimentin have been highly conserved. In contrast, the conspicuous lack of homology among the non- α -helical sequences of these IF subunits suggests that these peripherally located portions of the subunits contribute to the observed variations in the antigenicity and solubility properties which distinguish the types of IF. This structural arrangement conforms to and elaborates on our earlier proposals of IF structure^{2,4}.

We thank Drs Stuart Yuspa, Leslye Johnson and David Davies for encouragement, Dr Robert Jernigan for assistance with the computer analysis of the sequence data, and Mr William Idler and Ms Christina Cheng for technical assistance.

Received 23 November 1982; accepted 1 February 1983.

1. Lazarides, E. *Nature* **283**, 249–256 (1980).
2. Steinert, P. M. in *Electron Microscopy of Proteins* Vol. 1 (ed. Harris, J. R.) 125–166 (Academic, London, 1981).
3. Steinert, P. M., Zackroff, R. V., Vynardi-Whitman, M. & Goldman, R. D. *Meth. Cell Biol.* **24**, Pt A, 399–419 (1982).
4. Steinert, P. M., Idler, W. W. & Goldman, R. D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4534–4538 (1980).
5. Geisler, N., Kaufman, E. & Weber, K. *Cell* **30**, 277–286 (1982).
6. Steinert, P. M., Zimmerman, S. B., Starger, J. A. & Goldman, R. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6098–6101 (1978).
7. Steinert, P. M. & Yuspa, S. H. *Science* **200**, 1491–1493 (1978).
8. Steinert, P. M. *et al. Biochim. biophys. Acta* **577**, 11–21 (1979).
9. Roop, D. R., Hawley-Nelson, P., Cheng, C. K. & Yuspa, S. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 716–720 (1983).
10. Roop, D. R., Kristo, P., Stumph, W. E., Tsai, M.-J. & O'Malley, B. W. *Cell* **23**, 671–680 (1981).
11. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
12. Steinert, P. M., Wantz, M. L. & Idler, W. W. *Biochemistry* **21**, 177–183 (1982).
13. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211–245 (1974).
14. Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97–120 (1978).

15. McLachlan, A. D. *Biopolymers* **16**, 1271–1297 (1977).
16. Parry, D. A. D. in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* Vol. 1 (eds Parry, D. A. D. & Creamer, L. K.) 393–427 (Academic, London, 1979).
17. Sodek, J., Hodges, R. S., Smillie, L. B. & Jursasek, L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3800–3804 (1972).
18. McLachlan, A. D. & Stewart, M. J. *molec. Biol.* **98**, 293–304 (1975).
19. Parry, D. A. D. *J. molec. Biol.* **153**, 459–464 (1981).
20. Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. *J. molec. Biol.* **120**, 311–325 (1978).
21. Parry, D. A. D., Crewther, W. G., Fraser, R. D. B. & MacRae, T. P. *J. molec. Biol.* **113**, 449–454 (1977).
22. McLachlan, A. D. *J. molec. Biol.* **124**, 297–304 (1978).
23. Ahmadi, B., Boston, N. M., Dobb, M. G. & Speakman, P. T. in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* Vol. 2 (eds Parry, D. A. D. & Creamer, L. K.) 161–166 (Academic, London, 1979).
24. Steinert, P. M. *J. molec. Biol.* **123**, 49–70 (1978).
25. Crick, F. H. C. & Rich, A. *Nature* **176**, 780–782 (1955).
26. Hanukoglu, I. & Fuchs, E. *Cell* **31**, 243–252 (1982).
27. Geisler, N., Plessmann, U. & Weber, K. *Nature* **296**, 448–450 (1982).
28. Geisler, N. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4120–4123 (1981).
29. Tseng, S. C. G. *et al. Cell* **30**, 361–372 (1982).
30. Steven, A. S., Wall, J., Hainfeld, J. & Steinert, P. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3101–3105 (1982).

Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene

W. Charles Johnson, Charles P. Moran Jr & Richard Losick

Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

A developmentally regulated gene (*spoVG*) from the spore-forming bacterium *Bacillus subtilis* is expressed from two overlapping promoters, which direct transcription initiating from sites separated by 10 base pairs. Initiation of the upstream promoter is determined by an RNA polymerase sigma factor of molecular weight 37,000 (σ^{37}). We report the isolation of a 32,000-molecular weight species of sigma factor (σ^{32}), which exclusively dictates transcription initiation from the downstream promoter, and suggest a model for the way in which sigma-specific recognition sequences are intermeshed within the *spoVG* transcription initiation region.

A DIFFERENTIATING, Gram-positive bacterium, *Bacillus subtilis*, produces a dormant cell form known as the endospore in response to nutrient depletion. The initiation phase of this differentiation process is controlled by a class of regulatory genes known as the *spoO* loci; mutations in any one of these eight regulatory genes block development before the earliest morphological events of spore formation^{1,2}. To study the control of gene expression during spore formation, we previously cloned in *Escherichia coli* a *B. subtilis* gene whose transcription is developmentally regulated³. This gene, known as *spoVG*⁴, is activated at the onset of sporulation and its transcription is controlled by the products of at least five of the eight known *spoO* loci⁵. A curious feature of *spoVG* is that its transcription in sporulating cells is initiated from two closely spaced start points known as P1 and P2 (Fig. 1 and ref. 6). The upstream start point P1 is used *in vitro* by a modified form of *B. subtilis* RNA polymerase containing a 37,000 (37K) molecular weight (MW) species of *B. subtilis* sigma factor known as σ^{37} (refs 6–8). (The principal form of *B. subtilis* RNA polymerase found in vegetative cells contains a sigma of 55,000 MW (σ^{55}).) Here we report the isolation of a 32K species of sigma factor (σ^{32}), which exclusively dictates utilization of the downstream promoter P2. Promoter deletion analysis to be reported elsewhere (C. Banner, C.P.M. and R.L., unpublished results) demonstrates that polymerase containing σ^{37} ($E\sigma^{37}$) and polymerase containing σ^{32} ($E\sigma^{32}$) initiate transcription from promoter regions (P1 and P2) which overlap substantially. Based on nucleotide sequence homology to another promoter under dual σ^{37} and σ^{32} control, a model is proposed for the arrangement of sigma-specific recognition sequences within the region of promoter overlap.

Purification of RNA polymerase using the P2 promoter

Highly purified $E\sigma^{37}$ uses only the P1 promoter of *spoVG*, but we were previously able⁶ to detect in less purified preparations of σ^{37} -containing polymerase, an activity that initiated at the P2 startpoint. To isolate this P2 transcribing enzyme, RNA polymerase that had been partially purified by phase partitioning and gel filtration was applied to a DNA cellulose column and eluted with a linear gradient of KCl. Figure 2 shows the elution pattern of RNA polymerase as assayed with the synthetic template poly(dA-dT). To detect specific utilization of the *spoVG* promoters, RNA polymerase in fractions from across the gradient was challenged with a truncated *spoVG*-containing plasmid (pCB1291; ref. 4) that had been cleaved with the restriction endonuclease *EcoRI* at a site located 120 and 110 base pairs (bp), respectively, downstream from P1 and P2. RNA polymerase initiating at P1 generated a 'run-off' RNA of 120 bases while enzyme initiating at the downstream promoter produced a transcript of 110 bases. The inset in Fig. 2 shows that gradient elution from DNA cellulose partially resolved RNA polymerase initiating at the upstream startpoint from enzyme initiating at the downstream startpoint.

To isolate the form of polymerase responsible for transcription from P2, enzyme in fractions 18 to 21 were pooled and

Table 1 Conserved promoter sequences

Holoenzyme	'-35' region	'-10' region	
<i>B. subtilis</i> $E\sigma^{55}$	TTGACA	TATAAT	(9)
<i>B. subtilis</i> $E\sigma^{37}$	AGG-TT	GG-ATTG-T	(2)
<i>B. subtilis</i> $E\sigma^{32}$	AAATC	TA-TG-TT-TA	(2)
<i>B. subtilis</i> $E\sigma^{29}$	TT-AAA	CATATT	(4)
<i>B. subtilis</i> $E\sigma^{28}$	CTAAA	CCGATAT	(2)
Phage SPO1 $E\sigma^{sp28}$	T-AGGAGA--A	TTT-TTT	(6)
Phage SPO1 $E\sigma^{sp33-34}$	CGTTAGA	GATATT	(5)

The references for the conserved sequences are: $E\sigma^{55}$, ref. 17; $E\sigma^{37}$, refs 11, 12; $E\sigma^{32}$, this paper; $E\sigma^{29}$, ref. 20; $E\sigma^{28}$, ref. 19; $E\sigma^{sp28}$, ref. 18; and $E\sigma^{sp33-34}$, J. Pero, personal communication. The number of promoter sequences compared in each class is shown in parentheses.

the enzyme further purified by DEAE-Sephacel chromatography, followed by a second gradient elution from DNA cellulose. The run-off transcription assays of Fig. 3c show that polymerase in fractions from the final DNA cellulose chromatography step was completely free of P1 transcribing activity, generating exclusively a run-off transcript of 110 bases.

Isolation of a low abundance σ factor

The subunit composition of RNA polymerase utilizing the P2 promoter was examined by subjecting fractions of the final DNA cellulose gradient elution to electrophoresis in an SDS-polyacrylamide slab gel. Staining with Coomassie brilliant blue readily visualized the $\beta\beta'$ - and α -subunits of RNA polymerase as well as several contaminating polypeptides, but failed to reveal a polypeptide whose elution pattern coincided with enzyme capable of generating the 110-base run-off RNA (Fig. 3a). (Note that the P2 transcribing activity (Fig. 3c) eluted after the bulk of the RNA polymerase (Fig. 3a).) However, by treating the gel with the highly sensitive silver staining procedure of Oakley *et al.*⁹ we were able to detect a polypeptide whose elution profile largely coincided with that of P2 transcribing activity. This 32K polypeptide is shown in the inset of Fig. 3b, which displays silver-stained polypeptides in the region of the RNA polymerase α -subunit. (The correlation of P2 transcribing activity with 32K protein was readily apparent for fractions 20–26 but could not be clearly examined at the low-salt end of the gradient (fraction 18), which contained low levels of several proteins in the 30–34K range.)

To determine directly whether the predominant 32K species was responsible for dictating recognition of the P2 promoter, the peak fractions (22–26) of P2 transcribing activity were pooled and applied to an SDS-polyacrylamide slab gel. Slices were cut from the gel in the position of the 32K polypeptide (slice B) and from above (slice A) and below (slice C) the putative transcription factor. Protein was then eluted from the slices, renatured according to the procedure of Hager and Burgess¹⁰ and added to core RNA polymerase. The run-off transcription experiment of Fig. 4 shows that renatured protein from slice B (track b) but not from slices A or C (tracks a and c) conferred on core RNA polymerase the ability to utilize the



Fig. 1 Nucleotide sequence of the *spoVG* promoter region. The centre of the figure shows the nucleotide sequence of the non-transcribed strand of the *spoVG* promoter region⁶. The positions of bases are designated by their distance in nucleotides from a downstream *EcoRI* endonuclease restriction site. Bold bases identify the initiation codon for protein synthesis (position 85), the ribosome binding site (positions 93–103; see ref. 17) and an upstream A+T-rich box (positions 161–186), which deletion analysis (ref. 6 and C. Banner, C.P.M. and R.L., in preparation) has shown to be important in efficient utilization of both P1 and P2. Regions of homology of *spoVC* to the *spoVG* P1 and P2 promoters are shown at the top and bottom of the figure, respectively, with the large letters designating positions of congruence¹¹. Positions of close contact between $E\sigma^{37}$ and purines in the *spoVC* sequence¹² are indicated by asterisks (note that some of these positions represent purines on the transcribed strand, which is not shown).

P2 promoter. As a control, σ^{37} that had been gel-purified by the Hager and Burgess procedure exclusively directed utilization of the P1 promoter (Fig. 4e).

The 32K polypeptide, which we hereafter refer to as σ^{32} , is a low abundance polypeptide as isolated in association with RNA polymerase. Assuming $\sim 5,000$ RNA polymerase molecules per cell, we estimate the abundance of $E\sigma^{32}$ at ~ 30 molecules per cell (compared with ~ 500 σ^{37} molecules per cell). As we did not suffer substantial losses of P2 transcribing activity during the gradient elution steps, we do not believe

this low abundance reflects poor yield of our purification procedure. The cell may, however, contain additional molecules of σ^{32} that are not associated with core RNA polymerase.

Although the low abundance of σ^{32} precludes its direct detection without extensive purification, we were able to survey cells in different growth states for σ^{32} activity by using the run-off transcription assay. Like σ^{37} , σ^{32} was present in growing cells, in early sporulating cells and in stationary phase cells of the asporogenous mutant, SpoOA. (Indeed, SpoOA cells were used as the source of σ^{32} in the purification of Figs 2, 3.)

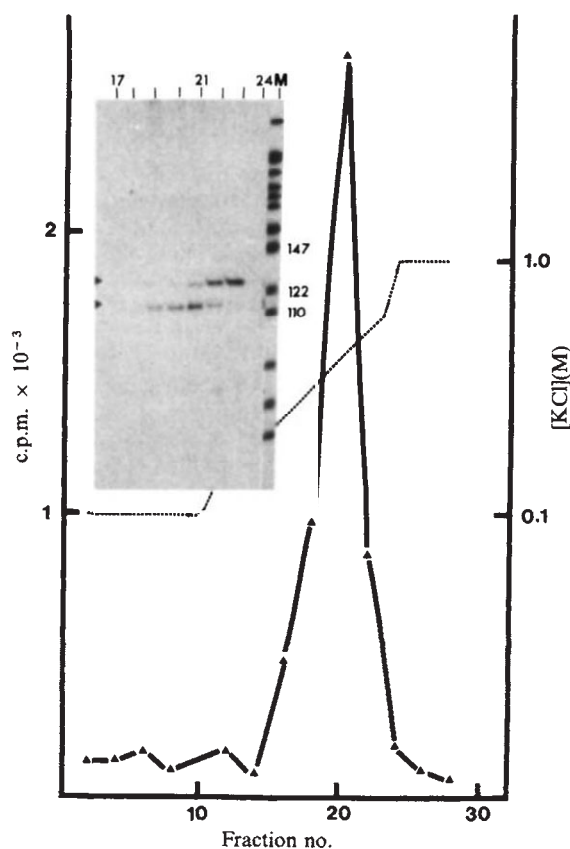
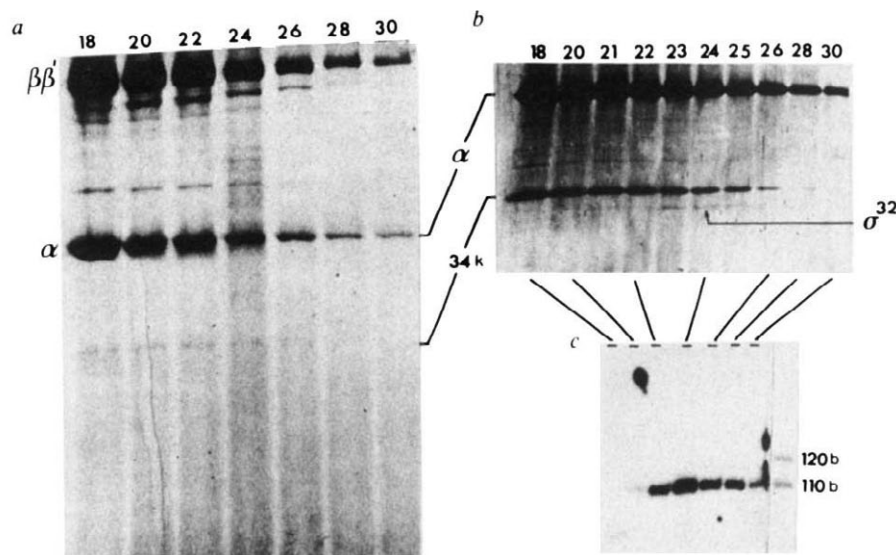


Fig. 2 Resolution of P1 and P2 transcribing activities by DNA cellulose chromatography. 100 g of *B. subtilis* cells (strain SpoOA-5NA) were collected 1 h after the end of logarithmic growth (T_1) in Difco sporulation medium²¹. Cells were disrupted by two passages through a French pressure cell (10,000 p.s.i.) and RNA polymerase was partially purified by phase partitioning²² and size fractionation on a Sephacryl S300 column (3 $\times$ 40 cm) equilibrated with buffer A (0.01 M Tris pH 8.0, 1 mM EDTA, 0.1 M KCl, 0.3 mM dithiothreitol (DTT), 0.3 mg ml⁻¹ phenylmethylsulphonyl fluoride (PMSF), 10% glycerol). Fractions from across the peak of RNA polymerase activity (determined by incorporation of ¹⁴C-AMP into RNA with poly(dA-dT) as template²²) were pooled and applied to a calf thymus DNA cellulose column (1 $\times$ 5 cm) and eluted with a 0.2–0.8 M KCl gradient (.....) in 50 ml of buffer C (0.05 M Tris pH 8.0, 1 mM EDTA, 0.3 mM DTT, 0.3 mg ml⁻¹ PMSF, 20% glycerol). Fractions (2 ml) were collected and samples (20 μ l) from each alternate fraction were assayed for RNA polymerase activity with poly(dA-dT) as template²² ($\blacktriangle$ – $\blacktriangle$); 2,300 c.p.m. corresponds to the incorporation of 1 nM of ¹⁴C-ATP into RNA in 10 min. Inset, run-off transcription. Samples (1 μ l) of DNA cellulose fractions 17–24 (each diluted 1:20) were incubated at 37 °C for 10 min with 2 μ g of *Eco*RI-cut pCB1291 DNA in 40- μ l reactions⁶. RNA synthesis was initiated by adding 0.5 mM each of ATP, GTP and UTP and 0.5 μ M (10 μ Ci [α -³²P]CTP. After 10 min at 37 °C, unlabelled CTP (1.5 mM) was added. After an additional 5 min of incubation the reaction was stopped. Radioactive RNA was purified from the reaction mixtures by ethanol precipitation and subjected to electrophoresis in a 6% polyacrylamide slab gel containing 7 M urea. Molecular weight markers (track labelled M) were ³²P-end-labelled fragments of *Hpa*II-cut pBR322 DNA which had been heat-denatured. The arrowheads on the left indicate the positions of the 120 (upper) and 110 (lower) run-off transcripts. Note that the run-off RNAs, whose sizes are known from nucleotide sequencing experiments, migrated slightly more slowly than the corresponding denatured DNA markers.

Fig. 3 Identification of σ^{32} . RNA polymerase in fractions 18–21 from the experiment of Fig. 2 were applied to a DEAE-Sephacel column (1 $\times$ 5 cm) and eluted with a linear salt gradient. Fractions of peak P2 transcribing activity were reapplied to a DNA cellulose column (1 $\times$ 5 cm) and eluted with 50 ml of a 0.35–0.85 M linear gradient of KCl in buffer C and 2-ml fractions were collected. 300 μ l of even-numbered fractions (18–30) were subjected to electrophoresis in a 12% polyacrylamide slab gel and stained with Coomassie brilliant blue (a) and to electrophoresis in a 10–20% gradient polyacrylamide slab gel



and stained with silver by the procedure of Oakley *et al.*⁹ (b). Molecular weight markers (not shown) covered the range 3,000–43,000 (protein molecular weight standards; BRL). In b, only silver-stained proteins in the MW range 25,000–45,000 are shown. c, 3 μ l of enzyme from even-numbered fractions (connected by solid lines to tracks in b) were used to transcribe *Eco*RI-cut pCB1291 DNA as described in the run-off transcription assay of Fig. 2. The 110- and 120-base run-off RNAs in the extreme right track were generated by a mixture of $E\sigma^{37}$ and $E\sigma^{32}$; molecular weight markers (not shown) were denatured ³²P-labelled fragments of *Hpa*II-cut pBR322 DNA.

σ^{37} and σ^{32} use the same startpoint in the *spoVC* promoter

The finding that two species of σ factor govern RNA synthesis from overlapping promoters prompted us to investigate whether other σ^{37} -controlled genes would be transcribed by σ^{32} -containing RNA polymerase. We used *spoVC*, a *B. subtilis* gene which is strongly transcribed by $E\sigma^{37}$ (refs 5, 8) and whose interaction with $E\sigma^{37}$ has been mapped in detail by 'DNase I footprinting'¹¹ and dimethyl sulphate methylation protection experiments¹². The run-off transcription experiment of Fig. 5 compares the ability of $E\sigma^{37}$ and $E\sigma^{32}$ to utilize *spoVC* and *spoVG* DNAs. The template for *spoVC* run-off transcription was plasmid DNA (pMI340) that had been cleaved at a unique *Bam*HI site (created *in vitro* by recombinant DNA techniques; M. Igo, unpublished work) located 155 bp downstream from the *spoVC* transcription startpoint. The template for *spoVG* transcription was *Eco*RI-cleaved pCB1291 DNA as described above. Figure 5 shows that both $E\sigma^{37}$ and $E\sigma^{32}$ were able to use the *spoVC* promoter and that the site of this transcription initiation was apparently identical for the two holoenzymes. Indeed, even when subjected to electrophoresis on a high resolution nucleotide sequencing gel (not shown), the run-off RNAs were indistinguishable in length. Utilization of the *spoVC* promoter by $E\sigma^{32}$ could not be attributed to contaminating σ^{37} because the purified $E\sigma^{32}$ enzyme was completely lacking in *spoVG* P1 transcribing activity (Fig. 5e). As additional evidence that $E\sigma^{32}$ could use the *spoVC* promoter, we have shown that enzyme reconstructed from core RNA polymerase and gel-purified σ^{32} in the experiment of Fig. 5 was also capable of promoting *spoVC* run-off transcription (data not shown).

Implications

The discovery of σ^{32} extends to seven the list of bacterial and phage-encoded sigma factors that govern the selection of promoters in *B. subtilis* (see Table 1). *B. subtilis* RNA polymerase is therefore emerging as a highly heterogeneous enzyme, capable of existing in multiple holoenzyme forms¹³. As each of these holoenzyme forms exhibits its own distinctive pattern of promoter recognition specificity, RNA polymerase heterogeneity may provide *B. subtilis* (and perhaps other species of Gram-positive bacteria) with a general mechanism for separately controlling the transcription of different classes of genes, including, in particular, developmental genes. In the case of *spoVG*, we observe a curious variation of this regulatory mechanism in which an interplay of two novel forms of holoenzyme ($E\sigma^{37}$ and $E\sigma^{32}$) governs the transcription of the same gene, albeit from separate, closely spaced startpoints.

The *spoVG* gene is under the complex control of a class of sporulation regulatory genes known as the *spoO* loci^{3,5}. It is unknown how these genes control *spoVG* expression or how their mode of action might be related to the dual promoters for $E\sigma^{37}$ and $E\sigma^{32}$. Nevertheless, it is interesting to note that dual promoters are not an uncommon feature of genes whose expression is under complex regulation. A pertinent example is the galactose-utilization (*gal*) operon of *E. coli*, which is transcribed from two promoters whose startpoints are separated by 5 bp (ref. 14). The two *gal* promoters are differentially controlled by the regulatory complex of cyclic AMP and cyclic AMP receptor protein (cAMP-CRP), which activates the downstream promoter and inhibits utilization of the upstream promoter. As the product of the first gene (*galE*) in the operon is synthesized several-fold more efficiently from the upstream transcript than from the downstream mRNA, differential utilization of the overlapping *gal* promoters in response to varying intracellular cyclic AMP levels has a magnified influence on the level of protein synthesis¹⁵. In this regard, it is tempting to suppose that competition between $E\sigma^{37}$ and $E\sigma^{32}$ for access to the *spoVG* transcription initiation region could modulate both the transcription of the *spoVG* gene and the translatability of its message. This possibility could be investigated by comparing the translational efficiencies of the P1 and P2 transcripts and

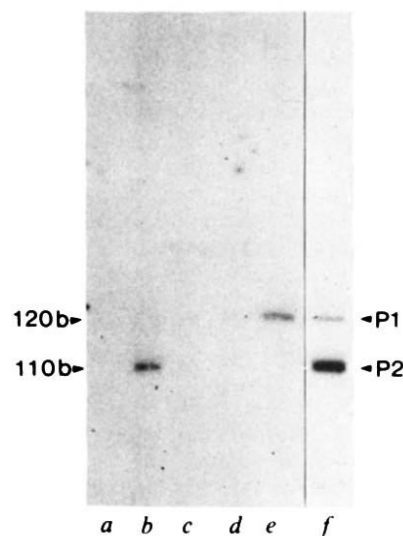


Fig. 4 Reconstruction of $E\sigma^{32}$ and $E\sigma^{37}$ from core RNA polymerase and gel-purified sigma. 5.0 μ g of $E\sigma^{32}$ -containing fractions 22–26 from the gradient elution of Fig. 3 and 54 μ g of the $E\sigma^{37}$ -containing fraction 22 from the gradient elution of Fig. 2 were subjected to electrophoresis in separate tracks of an SDS slab gel containing a 12–20% linear gradient of polyacrylamide. The gel was briefly stained with Coomassie brilliant blue and 0.5-cm slices were cut from above, below and at the position of σ^{32} or σ^{37} . These slices were then subjected to the protein elution and renaturation protocol of Hager and Burgess¹⁰. To reconstitute the holoenzymes, 15 μ l (representing 3% of the total) were incubated with 0.001 units of core RNA polymerase (given by W. G. Haldenwang) at 0 °C for 10 min. The reconstituted holoenzymes were then used to direct the transcription of *Eco*RI-cut pCB1291 and the 32 P-labelled products of RNA synthesis were subjected to electrophoresis in a 6% polyacrylamide gel containing 7 M urea. a, Slice above σ^{32} ; b, slice containing σ^{32} ; c, slice below σ^{32} ; d, slice above σ^{37} ; e, slice containing σ^{37} ; f, $E\sigma^{37}$ and $E\sigma^{32}$ mixture.

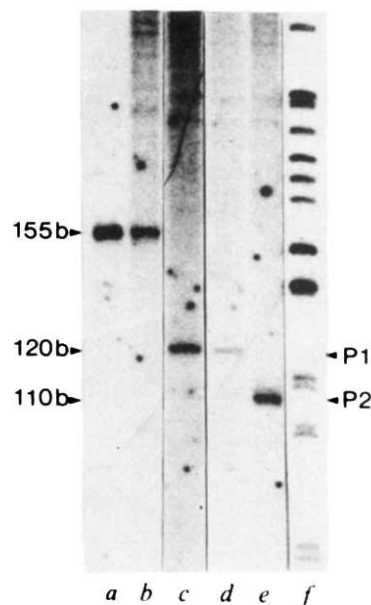


Fig. 5 Run-off transcription from *spoVC* DNA by $E\sigma^{32}$ and $E\sigma^{37}$. $E\sigma^{32}$ (30 ng) and $E\sigma^{37}$ (30 ng) were used to transcribe *Bam*HI-cut pMI340 (containing the *spoVC* promoter) DNA (2 μ g) or *Eco*RI-cut pCB1291 (containing the *spoVG* promoter) DNA (2 μ g). The 32 P-labelled run-off transcripts were subjected to electrophoresis in a 9% polyacrylamide gel containing 7 M urea. The transcriptions were of: a, *spoVC* DNA by $E\sigma^{37}$; b, *spoVC* DNA by $E\sigma^{32}$; c, d, *spoVG* DNA by $E\sigma^{37}$; and e, *spoVG* DNA by $E\sigma^{32}$. a, b, d and e were exposed to X-ray film for 20 h and c for 2 days. f, 32 P-labelled *Hpa*II fragments of pBR322.

by examining the regulatory consequences of mutations in the $E\sigma^{37}$ and $E\sigma^{32}$ promoters.

Transcription studies with deletion mutations of the *spoVG* promoter region (to be reported elsewhere; C. Banner, C.P.M. and R.L., in preparation) show that the promoters for $E\sigma^{37}$ and $E\sigma^{32}$ overlap substantially in the *spoVG* transcription initiation region, as was expected from the close spacing (10 bp apart) of the P1 and P2 startpoints. This suggests that this complex promoter region may consist of intermingled recognition sequences for the two *B. subtilis* holoenzyme forms.

In *E. coli*, which is thought to have a single form of holoenzyme, promoters generally exhibit two regions of approximate homology, which are centred ~35 (the '-35' region) and 10 (the '-10' region) bp upstream from the startpoint of transcription, and which are believed to determine recognition by RNA polymerase¹⁶. In *B. subtilis*, where RNA polymerase exists in multiple forms, different classes of promoters are recognized by different forms of RNA polymerase¹³. Nevertheless, each promoter class exhibits its own distinctive and conserved nucleotide sequences in regions centred ~35 and 10 bp upstream from the startpoint of transcription^{11,12,17-19}. Table 1 lists the consensus sequences for several of these *B. subtilis* chromosomal and phage promoter classes.

Figure 1 identifies sequences which we have previously suggested as being important for the recognition of P1 by $E\sigma^{37}$. This is based on the following: (1) Only two regions of homology can be identified between the *spoVC* promoter and the P1 promoter of *spoVG* and these are centred approximately at positions -35 and -10 (ref. 11). Note, however, that the -35 sequence for the *spoVG* P1 promoter is 3 bp closer to its startpoint than is the corresponding *spoVC* sequence to its startpoint. (2) These conserved regions lie within the region of *spoVC* protected from DNase by $E\sigma^{37}$, as shown by a DNA footprinting experiment¹¹, and within the boundaries of *spoVG* essential for promoter function, as shown by deletion analysis (ref. 6; C. Banner, C.P.M. and R.L., in preparation). (3) DNA methylation protection experiments revealed nine adenines and guanines in *spoVC* DNA that closely contact σ^{37} in open

promoter complexes¹², and all nine of these purines lie within or immediately adjacent to the two regions of homology to the *spoVG* P1 promoter (the position of these purines is indicated by asterisks in Fig. 1). (4) A series of promoters²⁰ that are weakly used by $E\sigma^{37}$ conform significantly in their -35 and -10 regions to the putative canonical recognition sequences ascribed to this enzyme.

Further comparison of the *spoVG* and *spoVC* promoters suggests the identity of the sequences that signal recognition by $E\sigma^{32}$ (as we have shown in Fig. 5 that *spoVC* is used by $E\sigma^{32}$ as well as by $E\sigma^{37}$). When aligned by their startpoints, the *spoVG* P2 and *spoVC* promoters exhibit two regions of significant homology upstream from their startpoints: an identical pentanucleotide sequence (AAATC) centred at position -30 and an 8 out of 11 position congruence (TA-TG-TT-TA) centred at position -7 (Fig. 1). (Additional homology (ATAGG) can be detected extending from the '-10' region into the startpoint region but at one base pair out of register.) The spacing between these homologous regions is almost identical in the two promoters (15 bp in the *spoVG* P2 promoter and 14 bp in the *spoVC* promoter). If the sequences AAATC and TA-TG-TT-TA do, in fact, signal promoter recognition by $E\sigma^{32}$, the *spoVG* transcription initiation region emerges as a mosaic of four separate sigma-specific recognition sequences arranged in the alternating order: σ^{37} '-35'; σ^{32} '-35'; σ^{37} '-10'; σ^{32} '-10' (Fig. 1). In *spoVC*, however, where $E\sigma^{37}$ and $E\sigma^{32}$ use common startpoints, the two -35 sequences abut one another and the two -10 sequences overlap substantially. In this case the two holoenzyme forms would be expected to interact with some of the same bases in their -10 regions. We plan to test our proposals for the arrangement of intermeshed recognition sequences in *spoVG* and *spoVC* by using *in vitro* mutagenesis to generate point mutations that separately or simultaneously impair promoter utilization by $E\sigma^{37}$ or $E\sigma^{32}$.

We thank J. Pero, P. Youngman and A. L. Sonenshein for helpful suggestions. This work was supported by a NIH postdoctoral fellowship (AI05918) to W.C.J. and a NIH grant (GM18568) to R.L.

Received 24 November 1982; accepted 17 February 1983.

- Hoch, J. A. in *Advances in Genetics* (ed. Caspari, E. W.) 67-98 (Academic, New York, 1976).
- Piggot, P. J. & Coote, J. G. *Bact. Rev.* **40**, 908-952 (1976).
- Segall, J. & Losick, R. *Cell* **11**, 751-761 (1977).
- Rosenbluh, A., Banner, C. D. B., Losick, R. & Fitz-James, P. C. *J. Bact.* **148**, 341-351 (1981).
- Ollington, J. F., Haldenwang, W. G., Huynh, T. V. & Losick, R. *J. Bact.* **147**, 432-442 (1981).
- Moran, C. P. Jr, Lang, N., Banner, C. D. B., Haldenwang, W. G. & Losick, R. *Cell* **25**, 783-791 (1981).
- Haldenwang, W. G. & Losick, R. *Nature* **282**, 256-260 (1979).
- Haldenwang, W. G. & Losick, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7000-7004 (1980).

- Oakley, B. A., Kirsch, D. R. & Morris, N. R. *Analyt. Biochem.* **105**, 361-363 (1980).
- Hager, D. & Burgess, R. *Analyt. Biochem.* **109**, 76-86 (1980).
- Moran, C. P. Jr, Lang, N. & Losick, R. *Nucleic Acids Res.* **9**, 5979-5990 (1981).
- Moran, C. P. Jr, Johnson, W. C. & Losick, R. *J. molec. Biol.* **162**, 709-713 (1982).
- Losick, R. & Pero, J. *Cell* **25**, 582-584 (1981).
- Musso, R. E., Di Lauro, R., Adhya, S. & deCrombrughe, B. *Cell* **12**, 847-854 (1977).
- Queen, C. & Rosenberg, M. *Cell* **25**, 241-249 (1981).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319-353 (1979).
- Moran, C. P. Jr et al. *Molec. gen. Genet.* **186**, 339-346 (1982).
- Lee, G. & Pero, J. *J. molec. Biol.* **152**, 247-265 (1981).
- Gilman, M. Z., Wiggs, J. L. & Chamberlain, M. J. *Nucleic Acids Res.* **9**, 5991-6000 (1981).
- Lang, N. thesis, Harvard Univ. (1982).
- Schaeffer, P., Millet, J. & Aubert, J. *Proc. natn. Acad. Sci. U.S.A.* **54**, 704-711 (1965).
- Shorenstein, R. G. & Losick, R. *J. biol. Chem.* **248**, 6163-6169 (1973).

LETTERS TO NATURE

Monopole heat

Michael S. Turner

Astronomy and Astrophysics Center, University of Chicago, Chicago, Illinois 60637, USA

The most stringent astrophysical limit on the flux of superheavy magnetic monopoles is based on the energy released when monopoles captured by a neutron star catalyse nucleon decay within the neutron star, and is: $F_G \leq 10^{-22} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$. This, and other astrophysical bounds, are bounds on the average flux of monopoles in the Galaxy, not on the local flux of monopoles. Dimopoulos *et al.*¹ have pointed out that the local flux could be many orders-of-magnitude larger than the average flux. Monopoles moving more slowly than about $3 \times 10^{-5} c$ ($10^{16} \text{ GeV m}^{-1}$) are stopped in the Earth. Inside the Earth they

catalyse nucleon decay, resulting in the release of heat. A bound on the local flux of monopoles moving more slowly than $3 \times 10^{-5} c$ ($10^{16} \text{ GeV m}^{-1}$), $F_{\oplus} \leq 10^{-21} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$, is derived here by requiring the energy released to be less than the measured heat flow at the surface of the Earth. Similar arguments are used to derive a bound on the flux of monopoles at Jupiter moving more slowly than $10^{-3} c$ ($10^{16} \text{ GeV m}^{-1}$), $F_J \leq 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$.

Polyakov² and 't Hooft³ have shown that when a semisimple group G breaks down to a group G' which contains a $U(1)$ factor there exist topologically-stable, classical solutions of the field equations whose long range fields are that of a magnetic pole of charge $n(2\pi/e) = n(69e)$ ($n = \text{integer}$). The energy associated with such a configuration, the mass of a monopole, is $O(M/\alpha)$ where M is the scale of spontaneous symmetry breaking, and α is the gauge coupling constant. For $SU(5) \rightarrow SU(3) \times SU(2) \times U(1)$, the monopole mass m is $O(10^{16} \text{ GeV}) \approx 10^{-8} \text{ g}$. Magnetic monopoles have other very interesting proper-

ties; they should catalyse nucleon decay⁴⁻⁶ and this process should proceed at a 'typical strong interaction rate'^{5,6}. The detection of a superheavy magnetic monopole would confirm the unification hypothesis, give cosmological evidence that the Universe was hotter than $0(10^{14})$ GeV, and have implications for astrophysics, if the flux exceeds bounds based on the survival of astrophysical magnetic fields⁷⁻¹¹. The candidate event reported by Cabrera^{12,13} corresponds to a 'flux' of $F \approx 6 \times 10^{-11} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$, which exceeds the bound based on the survival of the galactic magnetic field by some 5 orders-of-magnitude.

The most restrictive astrophysical bound is based on monopole catalysis of nucleon decay inside neutron stars¹⁴⁻¹⁶. Monopoles less massive than $0(10^{20})$ GeV which are incident on a neutron star lose sufficient energy to be stopped. Once inside they catalyse nucleon decay; the energy released is thermalized, and radiated in the form of photons and neutrinos. Requiring the photon flux to be consistent with observations results in the bound: $F_G \leq 10^{-22} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v/10^{-27} \text{ cm}^2)^{-1}$, where (σv) is the cross-section for catalysis times the monopole-nucleon relative velocity divided by c . However, this bound and other astrophysical bounds only apply to the average flux of monopoles in the Galaxy; the local monopole flux could be larger^{1,17}.

We now derive upper bounds on the flux of monopoles incident on the Earth with velocity $\leq 3 \times 10^{-5} c/m_{16}$ and on the flux of monopoles incident on Jupiter with velocity $\leq 10^{-3} c/m_{16}$ (monopole mass $m = m_{16} 10^{16}$ GeV). Monopoles moving this slowly lose sufficient energy to be stopped, and then catalyse nucleon decay, releasing heat. The limits, $F_\oplus \leq 10^{-21} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v/10^{-27} \text{ cm}^2)^{-1}$ and $F_J \leq 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v/10^{-27} \text{ cm}^2)^{-1}$, follow from requiring the rate of energy release from nucleon decay to be less than the measured amount of heat flowing out from the surface of the planet. [The issue is actually more complex. Global energy balance for the Earth or Jupiter involves three terms: heat flow at the surface; internal sources of heat; and the heat absorbed (or released) if the planetary temperature is increasing (or decreasing). For the Earth, radioactive decays of K, Th, and U account for essentially all of the heat flow at the surface, and so the assumption of constant Earth temperature would leave little room for additional sources. On the other hand, an additional heat source equivalent to the energy flow at the surface could be accommodated by allowing the Earth to heat up slowly. Rather than performing detailed calculations here, we will simply require the heat release from nucleon decay to be less than the measured heat flow at the surface of the planet.]

Monopoles moving through the Earth or Jupiter lose energy due to the Eddy currents they induce. Ahlen and Kinoshita¹⁸ have calculated that the rate of energy loss in the Earth is

$$dE/dx \approx 30 \text{ GeV cm}^2 \text{ g}^{-1} \beta \rho \quad (1)$$

where β is the monopole's velocity divided by c , and ρ is the density of the Earth. The fractional energy loss $\Delta E/E_0$ suffered by a monopole due to Eddy current losses is

$$\frac{\Delta E}{E_0} \approx 0.5 \times 10^{-4} I(\beta)/(m_{16} \beta_0^2) \quad (2)$$

where β_0 is the monopoles' initial velocity, and $\langle \beta \rangle$ is its average velocity while inside the Earth. The integral $I = \int \rho dx/(2R_\oplus \bar{\rho})$ depends on the impact parameter b of the monopole's trajectory. For $b = 0$, $I = 1.5$; for $b \leq 0.45 R_\oplus$, $I \geq 1.0$. The weighted average of I is: $\langle I \rangle = \int I 2\pi b db/\pi R_\oplus^2 \approx 3/4$. From equation (2) it follows that monopoles moving more slowly than about $\beta_0 \approx 3 \times 10^{-5} c/m_{16}$ will be stopped in the Earth due to Eddy current losses. [As the escape velocity from the Earth is $\sim 3 \times 10^{-5} c$ all monopoles will be moving at least this fast when they reach the surface. However, to become captured they need only lose their initial (far from the Earth) kinetic energy $= m\beta_0^2 c^2/2$, and so for all β_0 the criterion for capture is $\Delta E/E_0 \geq 1$. For monopoles with $\beta_0 \leq 3 \times 10^{-5}$ the value of $\langle \beta \rangle$ in equation (2) is $0(3 \times 10^{-5})$, rather than $0(\beta_0)$, making them slightly easier to capture.]

Ahlen and Kinoshita¹⁸ find that the rate of energy loss due to Eddy currents for monopoles moving through the Sun varies between 10 and $100 \text{ GeV cm}^2 \text{ g}^{-1} \beta \rho$, depending on the location in the Sun. Assuming a similar range for Jupiter, the fractional energy loss for a monopole passing through Jupiter is,

$$\frac{\Delta E}{E_0} \approx (0.4 - 4) \times 10^{-4} I(\beta)/(m_{16} \beta_0^2) \quad (3)$$

where as before $I = \int \rho dx/(2R_J \bar{\rho})$ depends on the impact parameter; for $b = 0$, $I \approx 3$. From equation (3), it follows that monopoles moving more slowly than about $\beta_0 \approx 10^{-3} c/m_{16}$ will be stopped in Jupiter due to Eddy current losses.

SU(5) monopoles also have colour magnetic fields, which are presumably screened at a distance of $\Lambda_{\text{QCD}}^{-1} \approx 0(1f)$, and because of this they can lose energy through 'ordinary' hadronic interactions as well as catalysis reactions. Assuming a cross-section of $0(10^{-27} \text{ cm}^2)$ and an energy loss per interaction of $0(100 \text{ MeV})$, the energy loss rate due to hadronic interactions is of the order of $10^{-4} \text{ GeV cm}^2 \text{ g}^{-1} \rho$, which might be comparable to the Eddy current losses for very slowly-moving monopoles. Drell *et al.*¹⁹ have calculated that monopoles passing through hydrogen or helium should lose energy at a rate of $0(5)$ times larger than in equation (1) due to transitions they induce between Zeeman-split atomic levels. However, as they point out, there may be threshold effects for velocities $\leq 2 \times 10^{-4} c$. The calculation has not yet been done for more complex atomic systems. Finally, if monopoles are also electrically-charged, for example, early during their passage through the Earth or Jupiter they might pick up a heavy nucleus (C. Goebel, unpublished data), then the energy loss rate can be significantly greater than the rate due to Eddy currents alone. If the nucleus picked up is Al, for example (S. Ahlen, personal communication), the rate of energy loss is about 30 times larger than in equation (1). Thus the fractional energy loss may be as much as $0(30)$ times our estimates in equations (2) and (3), so that monopoles moving as fast as $10^{-3} c/m_{16}$ may be stopped in the Earth and those moving as fast as $3 \times 10^{-2} c/m_{16}$ may be stopped in Jupiter.

Since its formation about 4,600 Myr ago the Earth should have captured N_M monopoles,

$$N_M = F_\oplus \times (4\pi R_\oplus^2) \times (\pi \text{ sr}) \times 4,600 \text{ Myr} \times f \\ \approx F_\oplus \times (2.4 \times 10^{36} \text{ cm}^2 \text{ sr s}) \times f \quad (4)$$

where f is the fraction of monopoles incident on the Earth with velocity $< 3 \times 10^{-5} c/m_{16}$, and $F_\oplus$ is the average local flux of monopoles during the past 4,600 Myr. Once captured, a monopole will catalyse nucleon decay at a rate $\Gamma \approx n(\sigma v)$, where n is the local number density of target baryons, and (σv) is the cross-section for catalysis of nucleon decay times relative velocity divided by c . The rate of energy release due to catalysed nucleon decay is then just

$$dE/dt = \int_{\text{earth}} n_M n(\sigma v) \langle E \rangle dV \quad (5)$$

where n_M is the local monopole number density, and $\langle E \rangle \approx$ is the average energy release. If the target baryons were just neutrons and protons (rather than nuclei), then n would be $6 \times 10^{23} (\rho/\text{g cm}^{-3})$ and $\langle E \rangle \approx 1 \text{ GeV}$. The number density of targets is actually $6 \times 10^{23} (\rho/\text{g cm}^{-3})/\langle A \rangle$, where $\langle A \rangle$ is the average atomic mass number. It is not clear what the energy release in an encounter between a nucleus of mass number A and a monopole is; it is certainly $\leq A \text{ GeV}$. If the captured monopoles were distributed uniformly throughout the Earth, then the integral in equation (5) could be done by replacing n_M and n by N_M and $\langle n \rangle$ respectively. It is very likely that monopoles will sink to the centre of the Earth where the density is $0(3)$ times greater than the average density of the Earth. Noting these uncertainties we integrate equation (5) with $n \approx 6 \times 10^{23} (\rho/\text{g cm}^{-3}) \approx 3.3 \times 10^{24}$ and $\langle E \rangle \approx 1 \text{ GeV}$, obtaining

$$dE/dt \approx 4 \times 10^{41} \text{ erg s}^{-1} \times F_\oplus \times f \times (\sigma v/10^{-27} \text{ cm}^2) \quad (6)$$

The factor $(\sigma v/10^{-27} \text{ cm}^2)$ represents the uncertainties in the catalysis cross-section. It could be much, much smaller than a

strong interaction cross-section⁴, or it could be strongly velocity dependent at low velocities because of strong interaction barriers (typical velocity of a nucleus at the centre of the Earth $\approx 10^{-5}c$)^{5,6}. In addition, unlike the neutron star case, the process may be complicated by interactions between the monopole and the atomic electrons (A. Goldhaber, personal communication).

The measured heat flow at the surface of the Earth is about $60 \text{ erg cm}^{-2} \text{ s}^{-1}$; integrating over the entire surface this is $3.0 \times 10^{20} \text{ erg s}^{-1}$ (ref. 20). If we require the heat generated by monopole-catalysed nucleon decay to be less than this amount, we obtain the bound,

$$F_{\oplus} \leq 7 \times 10^{-22} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \times f^{-1} \times (\sigma v / 10^{-27} \text{ cm}^2 \text{ s}^{-1})^{-1} \quad (7)$$

which is comparable to the bound on the average flux in the galaxy obtained by considering monopole catalysis in neutron stars¹⁴⁻¹⁶. However, it depends on the fraction f of monopoles striking the Earth with velocity $\leq 3 \times 10^{-5}c/m_{16}$.

For simplicity, let us assume that the distribution of monopole velocities is locally isotropic and uniform from $v = 0$ to $v = \alpha 10^{-3}c$. If the local flux of monopoles is primarily due to a cloud of orbiting monopoles, then $\alpha \approx 0.14$ since the escape velocity at 1 AU is $1.4 \times 10^{-4}c$. On the other hand, if the local flux of monopoles is primarily due to monopoles moving through the galaxy, then $\alpha \approx 0(1)$. The fraction of monopoles with velocity $\leq 3 \times 10^{-5}c/m_{16}$ relative to the Earth is about $(0.03/\alpha)^3 m_{16}^{-3}$. To obtain their fractional contribution to the local flux, f , we must multiply $(0.03/\alpha)^3 m_{16}^{-3}$ by their relative velocity divided by the average relative velocity ($\approx 0.03/\alpha m_{16}$), so that $f \approx (0.03/\alpha)^4 m_{16}^{-4}$. For the case of an orbiting cloud of monopoles $\alpha \approx 0.14$ and $f \approx 10^{-3} m_{16}^{-4}$, and so from equation (7) it follows that the local flux $F_{\oplus}$ must be $\leq 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v / 10^{-27} \text{ cm}^2 \text{ s}^{-1})^{-1} m_{16}^4$. If there is no local enhancement, then $\alpha \approx 0(1)$ and $f \approx 10^{-6} m_{16}^{-4}$, so that $F_{\oplus}$ must be $\leq 10^{-15} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v / 10^{-27} \text{ cm}^2 \text{ s}^{-1})^{-1} m_{16}^4$. Of course, in this case the neutron star bound is also applicable and is more restrictive. We note in passing that the fraction of monopoles moving slowly enough to be captured increases as $(dE/dx)^4$, so that if the rate of energy loss for monopoles moving through the Earth is larger than our estimate in equation (1), say by a factor of 3, then our bound becomes more restrictive by a factor of 10^2 .

The number of monopoles accreted by Jupiter since its formation is $N_M = F_J (2.9 \times 10^{38} \text{ cm}^2 \text{ sr s}) f$, where f is the fraction of monopoles incident on Jupiter with velocity $\leq 10^{-3}c/m_{16}$. The rate of energy release due to catalysed nucleon decay is given by an expression analogous to equation (5). Again, it is likely that monopoles will sink to the centre of Jupiter, where the density is 0(10) times greater than the average density. However, in light of the other uncertainties discussed above we have integrated equation (5) with $n = 6 \times 10^{23} (\bar{\rho}/g \text{ cm}^{-3}) = 8.4 \times 10^{23}$ and $\langle E \rangle \approx 1 \text{ GeV}$, obtaining

$$dE/dt \approx 1.1 \times 10^{43} \text{ erg s}^{-1} \times F_J \times f \times (\sigma v / 10^{-27} \text{ cm}^2 \text{ s}^{-1}) \quad (8)$$

Jupiter has a surface temperature of about 130 K, and radiates $\sim 1.0 \times 10^{25} \text{ erg s}^{-1}$ from its surface²⁰. About half of this energy is accounted for by radiant energy incident from the Sun. Requiring the heat generated by monopole catalysed nucleon decays to be $< 0.5 \times 10^{25} \text{ erg s}^{-1}$ results in the constraint,

$$F_J \leq 5 \times 10^{-19} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} (\sigma v / 10^{-27} \text{ cm}^2 \text{ s}^{-1})^{-1} f^{-1} \quad (9)$$

For 10^{16} GeV monopoles f should be 0(1), so that if there is a cloud of orbiting monopoles this limit is comparable to the Earth limit. If there is no local enhancement of monopoles, then this limit is significantly better than the Earth limit. For $m \geq 10^{16} \text{ GeV}$ f will be less than 0(1), and following our earlier analysis $f \approx 0(\alpha^{-4} m_{16}^{-4})$.

We conclude by mentioning that the heat flow at the surface of the Earth has previously been used to obtain a bound on the monopole-to-nucleon ratio in the Earth²¹, $n_M/n \leq 10^{-28}$ (here the 'heat source' considered was monopole-antimonopole annihilations), and on the proton lifetime, $\tau_p \geq 10^{20} \text{ yr}$ (unpublished data from C. Townes, V. L. Fitch, and M.S.T.).

This work was supported in part by the DOE through contract AC02-80ER10773A003 at Chicago. I thank Edward Anders and Frank Richter for useful discussions.

Received 2 November 1982; accepted 7 February 1983.

1. Dimopoulos, S., Glashow, S. L., Purcell, E. M. & Wilczek, F. *Nature* **298**, 824 (1982).
2. Polyakov, A. M. *Pis'ma Zh. Eksp. Teor. Fiz.* **20**, 430 (1974); *JETP Lett.* **20**, 194 (1974).
3. 't Hooft, G. *Nucl. Phys.* **B79**, 276 (1974); **B105**, 538 (1976).
4. Wilczek, F. *Phys. Rev. Lett.* **48**, 1146 (1982).
5. Callan, C. G. *Phys. Rev. D* **25**, 2141 (1982); *Phys. Rev. D* **26**, 2058 (1983).
6. Rubakov, V. A. *Pis'ma Zh. Eksp. Teor. Fiz.* **33**, 658 (1981); *JETP Lett.* **33**, 644 (1981); *Nucl. Phys.* **203B**, 311 (1982).
7. Parker, E. N. *Astrophys. J.* **160**, 383 (1970).
8. Lazarides, G., Shafi, Q. & Walsh, T. F. *Phys. Lett.* **100B**, 21 (1981).
9. Bludman, S. A. & Ruderman, M. A. *Phys. Rev. Lett.* **36**, 840 (1976).
10. Turner, M. S., Parker, E. N. & Bogdan, T. J. *Phys. Rev. D* **26**, 1296 (1982).
11. Rephaeli, Y. & Turner, M. S. *Phys. Lett.* **121B**, 115 (1983).
12. Cabrera, B. *Phys. Rev. Lett.* **48**, 1378 (1982).
13. Racine, W. *Monopole Meet.* (1982).
14. Dimopoulos, S., Preskill, J. & Wilczek, F. *Phys. Lett.* **119B**, 320 (1982).
15. Kolb, E. W., Colgate, S. A. & Harvey, J. A. *Phys. Rev. Lett.* **49**, 1373 (1982).
16. Bais, F. A., Ellis, J., Nanopoulos, D. V. & Olive, K. A. *CERN Preprint TH 3383* (1982).
17. Freese, K. & Turner, M. S. *Phys. Lett. B* (in the press).
18. Ahlen, S. & Kinoshita, K. *Phys. Rev. D* **26**, 2347 (1982); Martim'yanov, V. P. & Khakimov, S. Kh. *JETP* **35**, 20 (1972).
19. Drell, S., Kroll, N., Mueller, M., Parke, S. & Ruderman, M. *Phys. Rev. Lett.* **50**, 644 (1983).
20. Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1976).
21. Carrigan, R. A. Jr *Nature* **288**, 348 (1980).

Structure of the local interstellar medium and the line of sight to α Oph

Francesco Paresce*

Arcetri Astrophysical Observatory, Florence, Italy

The precise physical state and spatial distribution of interstellar matter within $\approx 100 \text{ pc}$ of the Sun has been the subject of intense speculation because of the implications its knowledge would have on interstellar medium theory^{1,2}, soft X-ray and EUV transparency^{3,4} and, possibly, for very nearby objects, on solar and terrestrial evolution^{5,6}. The general consensus holds that the local interstellar medium out to $\approx 100 \text{ pc}$ from the Sun is exceptionally clear of any condensation of density $\geq 0.1 \text{ cm}^{-3}$. Presumably, this is due to the effect of a supernova or stellar wind blast wave sweeping through the local interstellar medium in the relatively recent past^{7,8}. I point out here that a surprising number of independent lines of observational evidence unambiguously imply the existence of a large, moderately dense ($n_H \geq 1 \text{ cm}^{-3}$), cool ($T \approx 70 \text{ K}$) cloud of interstellar gas lying between the Sun and α Oph, an A5V star located $16.7 \pm 0.8 \text{ pc}$ away⁹ in the direction $l = 32^\circ$, $b = 23^\circ$.

Vidal-Madjar *et al.*¹⁰ and Bruston *et al.*¹¹ have suggested, on the basis of indirect arguments concerning the possible spatial inhomogeneities of the local deuterium abundance, that a dense interstellar cloud might exist within 0.01–2 pc of the Sun in the general direction of the Scorpius–Centaurus association. The large experimental and theoretical uncertainties and the lack of data on this constituent have led some to question, however, the need for any large density discontinuities in the local interstellar medium¹². This conclusion seems to be borne out by the results of an interstellar polarization survey¹³ carried out recently to $\approx 35 \text{ pc}$ from the Sun. This survey failed to find any direct evidence of significant dust concentration in the portion of the local interstellar medium observed so far although there may be some weak evidence for a relatively nearby patch of gas below the plane in the general direction of the galactic centre. Moreover, soft X-ray all-sky background observations in the B band (effective photon energy $\approx 150 \text{ eV}$) for which unit optical depth is reached at neutral hydrogen column densities $N(\text{H I}) \approx 5\text{--}10 \times 10^{19} \text{ cm}^{-2}$ also show no evidence for this or any other dense cloud within $\approx 100 \text{ pc}$ of the Sun¹⁴. Incidentally, this result may have a serious discrepancy with

*Present address: Space Telescope Science Institute, Baltimore, Maryland 21218, USA.

the predictions of theories describing the state of the interstellar medium during or after the passage of a supernova generated shock wave in an inhomogeneous substrate¹ and with the results of precise Stromgren photometry of nearby A and F stars¹⁵. The latter imply the existence of numerous clouds in the interstellar medium beyond ≈ 50 pc of average neutral hydrogen density $n_{\text{H}} \approx 30 \text{ cm}^{-3}$ and diameters of ≈ 4 pc. The opposite view of an adiabatic blast wave propagating in a fairly cloud-free interstellar medium has been described in detail by Cox and Anderson².

The negative results just quoted may, however, be due to very small dense or very large diffuse clouds escaping detection due to the very low spatial resolutions and highly incomplete sky coverage of these surveys and the marginal absorption of photons in the soft X-ray region by all but the densest clouds. Because of the serious implications on interstellar medium theory a confirmation of these conclusions would entail, considerable effort has been expended on the task of probing the structure of the local interstellar medium by means of high spectral resolution observations of nearby stars and solar resonance backscatter techniques^{16,17}. All these results seem to converge on the remarkable conclusion that the Sun is immersed in a warm ($T \approx 10^4$ K), partially ionized ($n_{\text{H}}/n_{\text{H}^+} \approx 0.5$), low density ($n_{\text{H}} \approx 0.05\text{--}0.1 \text{ cm}^{-3}$) gas extending out to ≤ 5 pc and that no other denser material is present anywhere within a ≈ 50 pc radius of the Sun.

A surprising aspect of these investigations is that they seem to have completely overlooked the implications of the numerous but scattered observations of α Oph (Rasalhague) that I have summarized for convenience in Table 1 which lists the five indirect ways in which $N(\text{H I})$ to this star may be estimated. In the following, I assume the cosmic abundances of the elements relative to hydrogen, A_{x} , tabulated by Allen¹⁸.

The first estimate in Table 1 is based on the observed Ti II, 3,384 Å line absorption profile in the α Oph spectrum and the observed range of interstellar medium titanium depletions relative to hydrogen reported by Stokes¹⁹. The depletions range from $\log \delta(\text{Ti}) = \log [N(\text{Ti II})/A_{\text{Ti}}N(\text{H})] \approx -1.3$ characteristic of stars with wide shallow absorption profiles to $\log \delta(\text{Ti}) = -2.3$ for stars with narrow deep profiles. The α Oph 3,384 Å absorption feature is typical of the latter category and is in dramatic contrast to the very shallow profile of λ Sco, a B1 star located in the galactic plane $\approx 50^\circ$ away from α Oph and at a distance of 120 pc from the Sun. The α Oph profile is dominated by a component centred at -26 km s^{-1} heliocentric velocity. Adopting these values and ignoring H_2 contributions to $N(\text{H})$ in such a short path, I obtain $19.4 \leq \log N(\text{H I}) \leq 20.4$ with the most likely value being 20.1 for the average $\log \delta(\text{Ti}) = -2$ for cloud-like absorption profiles.

The second method relies on the measured Na I column density to α Oph obtained by Hobbs²⁰ using the D1 line at 5,896 Å and on the expression $\log N(\text{H}) = [\log N(\text{Na I}) + 31.3]/2.11$ suggested by Stokes as the least squares solution for the observed $N(\text{Na I})$ against $N(\text{H})$ relation in the solar neighbourhood. Use of these values in this case yields $\log N(\text{H}) = 19.8 \pm 0.5$ judging from the scatter in the original data. The third technique uses the Fe II, 2,599 Å line absorption and density determination to α Oph reported by Frisch²¹. The measured value of $\log N(\text{Fe II})$ corresponds to $\log N(\text{H I}) = 20.1 \pm 0.5$ assuming an average interstellar Fe/H abundance ratio of 4.5×10^{-7} and the scatter about this value inferred from data reported by Savage and Bohlin²². The observed Ca II abundance is also consistent with all the quoted values. Use of the solar abundances²³ or an alternate cosmic abundance²⁴ increases all these numbers by ≈ 0.2 to 0.4 in the log.

The fourth method employs the Mg II h and k chromospheric emission line profiles at 2,813 and 2,834 Å that can be used to yield the $N(\text{Mg II})$ to the star. Although the Mg depletion $\delta(\text{Mg})$ in the direction of α Oph is unknown, we can assume it is the same as that of the absorbing gas in front of α Cen ($r = 1.35$ pc), λ And ($r = 23$ pc) and HR1099 ($r = 33$ pc) for which there exist simultaneous $N(\text{H I})$ and $N(\text{Mg II})$ column

Table 1 log of column density (atoms cm^{-2}) to α Oph

H I	Ti II	Ca II	Na I	Fe II	Mg II	Refs
19.4– 20.4	11.22 ± 0.08	11.18				19, 20
19.8 20.1			10.48	13.78 ± 0.15		20 21
19.6– 20.6 ≤19.7					15.01 ± 0.73	21 32

densities^{12,25–27}. These observations yield $\delta(\text{Mg}) = 0.008\text{--}0.16$, a range encompassing the average interstellar medium $\delta(\text{Mg})$ of ≈ 0.1 . Adopting this range, I obtain $19.6 \leq \log N(\text{H I}) \leq 20.6$ with most of the uncertainty due to the large observational error. Finally, the total velocity integrated $N(\text{H I})$ from the Berkeley 21-cm radio survey in the direction of α Oph is $\log N(\text{H I}) = 20.9$ while the amount of gas in the -20 to -30 km s^{-1} velocity range corresponding to the Ti II absorption feature amounts to $\log N(\text{H I}) \approx 19.7$.

Thus, although one might expect that any one of these five techniques might be subject to considerable systematic error, it is remarkable that they all give essentially the same answer to within the quoted uncertainties that is $19.4 \leq \log N(\text{H I}) \leq 20.4$. The lower bound to $N(\text{H I})$ implies an average volume density to the gas to α Oph of $n_{\text{H}} \geq 0.5 \text{ cm}^{-3}$, the lower limit obtained assuming the gas is uniformly distributed along the sun– α Oph path. This limit is already at least a factor of seven larger than the density expected on the basis of column densities to all the other stars in the solar vicinity analysed so far¹⁷ ($\approx 2 \times 10^{17} \text{ cm}^2 \text{ pc}^{-1}$) and may be as high as 5 cm^{-3} if $\log N(\text{H I}) = 20.4$.

An important additional clue to the nature of the local interstellar medium and the validity of the hypothesis I have just discussed, is represented by the peculiar -26 km s^{-1} heliocentric velocity of the dominant Ti II absorption feature in the α Oph spectrum. This velocity corresponds very closely to the projection along the line of sight to α Oph of the bulk velocity of a gas moving at -28 km s^{-1} from the direction $l = 25^\circ$, $b = 10^\circ$ deduced from observations of a large sample of stars in the local interstellar medium²⁸. Thus, the absorbing cloud to α Oph must be taking part in the general coherent motion of the local interstellar medium including that of the very local gas sampled within the Solar System by solar backscatter techniques.

This fact has three important implications. First, it makes it very unlikely that the observed α Oph spectral features listed in Table 1 are attributable to circumstellar rather than interstellar matter since the stellar radial velocity²⁹ ranges from $+6$ to $+14 \text{ km s}^{-1}$. Second, it confirms the basic premise used in generating $N(\text{H I})$ from the observed $N(\text{Ti II})$, $N(\text{Na I})$, $N(\text{Mg II})$, and $N(\text{Fe II})$, that the appropriate element depletions are typical of the general local interstellar medium gas moving at the same speed³⁰ and are the same as those found for the cloud in which the Sun is immersed and others around it. Thus, for example, the $N(\text{Mg})/N(\text{H})$ ratios measured for α Cen, λ And, HR1099 can be considered straightforwardly applicable to the α Oph cloud. Third, the proximity of the α Oph cloud speed with that of the local interstellar medium and with the general direction of the north polar spur (NPS), naturally suggests that it may be part of the expanding bubble of gas originating somewhere from the direction of the Sco–Cen association⁷. In fact, the gas in front of α Oph may be simply part of a much larger complex extending all the way back into the NPS with uniformly increasing temperature along this line of sight. The local gas in which the Sun is immersed would represent the warm ionized edge of the advancing front that is elongated in the direction of the NPS. This gas would be heated by conduction from the hot ($10^5\text{--}10^6$ K) interstellar medium and ionized by the soft X-ray/EUV background as envisioned in the McKee and Ostriker theory¹.

The Na/H ratio in the α Oph cloud ranges from 10^{-10} to 10^{-9} according to the results listed in Table 1, with a probable value of 4×10^{-10} . This value corresponds to the typical 'weak Na I absorption' category discussed by Hobbs³¹ who has shown that this ratio implies an absorbing gas that is warmer and/or less dense than the standard interstellar cloud material. Specifically, the observed α Oph Na/H ratio implies, assuming $T \approx 70$ K, that the gas is distributed over most of the line of sight with densities of the order of $n_H \leq 2-3 \text{ cm}^{-3}$ and low fractional ionization ($n_e \leq 10^{-3} \text{ cm}^{-3}$). The low recombination rate required could also be obtained by increasing T at the expense of n_H but the latter would then conflict with the lower limit $n_H \geq 0.5 \text{ cm}^{-3}$ determined earlier. Consequently, it seems most likely that the bulk of the absorbing gas towards α Oph is cool and relatively dense with probable positive density and negative temperature gradients along the line of sight.

These conclusions are consistent, broadly speaking, with those obtained by Vidal-Madjar *et al.*¹⁰ using more indirect arguments and may support Tinbergen's¹³ finding that there may be a nearby diffuse patch of dust in this general area. Taken together, these results may imply the existence of a large, perhaps inhomogeneous cloud extending over a large region of the sky in the direction of the galactic centre.

Of more than passing interest in this particular context is the behaviour of two relatively nearby stars: δ Cyg ($l = 79^\circ$, $b = 10^\circ$, $r = 48$ pc) and σ Sgr ($l = 10^\circ$, $b = -12^\circ$, $r = 57$ pc). Both these stars exhibit weaker Ti II, 3,384 Å components than α Oph by a factor of ≈ 4 but both are again precisely at the expected local interstellar medium coherent velocity for their directions²⁸. The identical shapes and velocities but different equivalent widths suggest that the same gas may be responsible for the absorption seen in all three stars but the lines of sight to δ Cyg and σ Sgr just graze the cloud. If so, the cloud occupies $\approx 70^\circ$ in the sky in the first quadrant, another weak argument for its relative proximity. The effect of this cloud on the soft X-ray background may be difficult to detect because of its diffuseness and large angular extent. Background measurements in the less penetrating EUV ($\lambda \approx 200$ Å) would be better suited to the task.

This work was carried out under the auspices of a NATO visiting professorship at Arcetri. Professor Franco Pacini's hospitality and encouragement are gratefully acknowledged. I also thank Professor Harold Weaver for insights and Dialta Paresce for editorial assistance.

Received 4 October 1982; accepted 23 February 1983.

- McKee, C. & Ostriker, J. P. *Astrophys. J.* **218**, 148 (1977).
- Cox, D. P. & Anderson, P. R. *Astrophys. J.* **253**, 268 (1981).
- Crutcher, R., Paresce, F., Bowyer, S. & Lampton, M. *Astrophys. J.* **187**, 497 (1974).
- Paresce, F. in *Astrophysics from Space* Vol. 81 (eds Bernacca, P. L. & Ruffini, R.), 243 (Reidel, Dordrecht, 1980).
- Begelman, M. C. & Rees, M. J. *Nature* **261**, 298 (1976).
- McKay, C. P. & Thomas, G. E. *Geophys. Res. Lett.* **5**, 215 (1978).
- Weaver, H. F. *IAU Symp. No. 84*, 295 (1978).
- Davelaar, J., Bleeker, J. A. M. & Deerenberg, A. J. M. *Astr. Astrophys.* **92**, 231 (1980).
- Gliese, W. *Catalogue of Nearby Stars*, 22 (Veroff, Heidelberg, 1969).
- Vidal-Madjar, A., Laurent, C., Bruston, P. & Audouze, J. *Astrophys. J.* **223**, 589 (1978).
- Bruston, P., Audouze, J., Vidal-Madjar, A. & Laurent, C. *Astrophys. J.* **243**, 161 (1981).
- McIntock, W., Henry, R. C., Linsky, J. L. & Moos, H. W. *Astrophys. J.* **225**, 465 (1978).
- Tinbergen, J. *Astr. Astrophys.* **105**, 53 (1982).
- Fried, P. M., Nousek, J. A., Sanders, W. T. & Kraushaar, W. L. *Astrophys. J.* **242**, 987 (1980).
- Knude, J. *Astr. Astrophys. Suppl.* **38**, 407 (1979).
- Paresce, F. *IAU Rep. Astr.* **18A**, 651 (1982).
- Cash, W., Bowyer, S. & Lampton, M. *Astr. Astrophys.* **80**, 67 (1979).
- Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
- Stokes, G. M. *Astrophys. J. Suppl.* **36**, 115 (1978).
- Hobbs, L. M. *Astrophys. J.* **191**, 381 (1974).
- Frisch, P. C. *Nature* **293**, 377 (1981).
- Savage, B. D. & Bohlin, R. C. *Astrophys. J.* **299**, 136 (1979).
- Withbroe, G. D. in *Menzel Symposium* (NBS SP-353, U.S. Government Printing Office, 1971).
- Cameron, A. G. W. in *Origin and Distribution of the Elements* (ed. Ahrens, T.) (Pergamon, Oxford, 1968).
- Oegerle, W. R., Kondo, Y., Stencel, R. E. & Weiler, E. J. *Astrophys. J.* **252**, 302 (1982).
- Anderson, R. C. & Weiler, E. J. *Astrophys. J.* **224**, 143 (1978).
- Baliunas, S. L. & Dupree, A. K. *Astrophys. J.* **227**, 870 (1979).
- Crutcher, R. M. *Astrophys. J.* **254**, 82 (1982).
- Wilson, R. E. *General Catalogue of Stellar Radial Velocities* (Carnegie Institution, Washington DC, 1952).
- Shull, J. M., York, D. G. & Hobbs, L. M. *Astrophys. J. Lett.* **211**, L139, 211 (1977).
- Hobbs, L. M. *Astrophys. J.* **222**, 491 (1978).
- Heiles, C. *Astr. Astrophys. Suppl.* **20**, 37 (1975).

Has rapid solar core rotation been observed?

B. Nyborg Andersen & P. Maltby

Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo 3, Norway

Internal rotation and gravitational quadrupole moment of the Sun are of interest to solar physics, the study of stellar structure and to investigations related to the test of gravitational theories. High precision measurements of fluctuations in the limb darkening function and in the spectral line shifts have raised the possibility that the interior of the Sun may be studied more directly than had previously been possible. Recently, Claverie *et al.*¹ argued that their detection of a 13.1 ± 0.2 day velocity signal give further experimental evidence that the solar core is rotating more rapidly than the observable surface. We show here that the phase as well as the magnitude of the observed signal amplitude may be predicted without any rapid core rotation by taking into account the presence of sunspots and their contribution to the spectral line profile as integrated over the disk of the Sun. Hence, we conclude that the existence of a 13.1-day apparently periodic velocity signal with amplitude 6.5 m s^{-1} during the 88 days observing period cannot be taken as evidence for a rapidly rotating solar core.

The observations¹ give the line-of-sight velocity of the integrated solar surface by means of Doppler shift measurement of the 769.9-nm line of neutral potassium, using the optical resonance scattering technique. The reported velocity is based on observations between June and August 1981 on two excellent island sites. The 13-day periodic velocity signal is most apparent in the data from Izana, whereas the data from Haleakala show the same characteristics, but with a larger scatter. We note, however, that large deviations from the 13-day periodic behaviour do occur in the first 30 days of the 88 observing days at Izana.

As the instrument has very high spectral resolution, but a low angular resolution the contribution from different parts of the rotating solar surface to the velocity signal will be weighted² depending mainly on heliographic coordinates and line-of-sight velocity. As far as we know, the possible influence of sunspots present on the solar disk has not been taken into account in these calculations. But by using calculations similar in nature to those of Brookes *et al.*², with the effects of sunspots included, an apparent velocity signal similar to that observed is produced.

Let us first consider the effect for a hypothetical sunspot with area 10^{-3} of the solar hemisphere positioned at the Sun's equator and with an intensity 0.5 times that of the photosphere throughout the spectral line. The presence of a spot on the solar disk means that compared with the 'unspotted' disk the contribution from the spot area is less. Before the spot's central

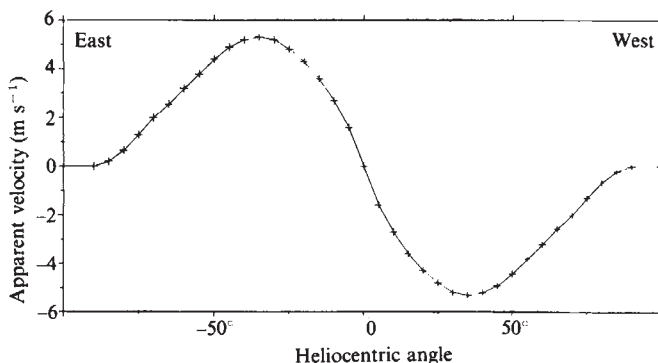


Fig. 1 The velocity signal as caused by a hypothetical sunspot with area 10^{-3} of the solar hemisphere and intensity equal to 0.5 times that of the photosphere is plotted against heliocentric angle.

meridian passage (CMP) this corresponds to less contribution from an area that has a line-of-sight velocity directed towards the observer. In the integrated line profile this corresponds to a slight displacement in the opposite direction, that is to the red. After CMP the line-of-sight velocity of the spot region is directed away from the observer, whereas the corresponding velocity signal is negative. In Fig. 1 the velocity signal caused by this hypothetical sunspot is plotted against heliocentric angle.

To test our model we have calculated the expected line profile for the whole Sun by adding the contribution to the line profile from each part of the visible hemisphere. The calculations were carried out by dividing the visible hemisphere into units of 20×10^{-6} ; that is approximately 5×10^4 equal areas. For each undisturbed area the continuum intensity is known from observations³ of the centre-to-limb variation. The continuum intensity ratios umbra/photosphere and penumbra/photosphere are observed^{4,5} to be typically 0.14 and 0.81, respectively for this wavelength region. As to the potassium 769.9-nm line itself we have used the limb darkening $I(\mu)/I(1) = 0.57 + 0.43 \mu$ as given by Brookes *et al.*².

Less information is available on the line profile in sunspots. However, J. W. Brault and O. Engvold (personal communication) have given us the following information based on unpublished spectra of large sunspots and the photosphere with the Kitt Peak National Observatory's Fourier-transform spectrometer. Comparing the undisturbed photosphere to the sunspot umbra the intensity ratio line centre/continuum decreases from 0.15 to 0.11. Correspondingly the widths (FWHM) of the line are 16.0 and 55.0 pm for the photosphere and the umbra, respectively. Assuming that the ratio between the penumbra and umbra radii is 2.4 for all sunspots, we estimated an average central intensity of 0.14 and an average line width of 23.5 pm for the sunspot region. The possible contribution from the broad wings in the sunspot line profile was included by adding a gaussian profile with weight 0.1 and width 70 pm to the 23.5 pm profile. However, the broad wings have only minor influence on the results. The line profile calculations also require that we know the line-of-sight velocity caused by the solar differential rotation⁶ for each unit area. As the resonance line of K I is formed relatively high in the solar atmosphere, where the effects of convective motions are small, we have carried out the calculations without taking the corresponding line shift (limb effect) into account⁷. Inside the sunspot umbra the limb effect appears to be absent⁸, whereas the Evershed flow will be present in the penumbra. As a first approximation we have neglected all motions except the solar rotation in calculating the line-of-sight velocity. In this simple model we have neglected projection effects like the tilt of the solar rotation axis and the effect of faculae. Finally we need to know the sunspot size and heliographic coordinates. The sunspot data (NOAA) for all sunspots with apparent area $\geq 20 \times 10^{-6}$ of the solar hemisphere are used to calculate the expected velocity signal for the Sun as a whole for each day during the observing period. The results are plotted in Fig. 2 together with the velocity signal as observed by Claverie *et al.*¹.

It appears from Fig. 2 that taking into account the passage of sunspots over the solar disk the calculated signal shows good agreement with the observed velocity signal at Izana, the correlation coefficient, r , is 0.64 (+0.06, -0.07). For the period 28 June to 22 August (the time span used by Claverie *et al.*¹ to detect the periodic signal) the corresponding correlation coefficient is 0.72; hence, our model shows higher correlation with the Izana data than do the supporting Haleakala observations ($r=0.40$). We note that the transitions from negative to positive signals and vice versa occur on practically the same dates as the observed values. The probability of the two signals being uncorrelated is $< 7 \times 10^{-9}$. It is of particular interest to consider the first 30 days of the Izana data that do not fit a 13-day period, $r = -0.26 (+0.22, -0.19)$. Our calculated signal, on the other hand, shows a correlation coefficient, $r = 0.50 (+0.15, -0.16)$ with the observations for this period.

Figure 2 shows that our model may mimic a periodic velocity signal for a limited number of days. It is to be expected (see Fig. 1) that the frequency power spectrum shows a peak close to half that of the synodic rotation period of the Sun. Comparing the calculated and observed frequency spectra we find good agreement with periods of 12.8 and 13.1 days, respectively. The calculated frequency spectrum depends mainly on the distribution of sunspots in longitude as well as on their area versus time variations. It is equally important that the model may also account for time spans without periodicity in the signal.

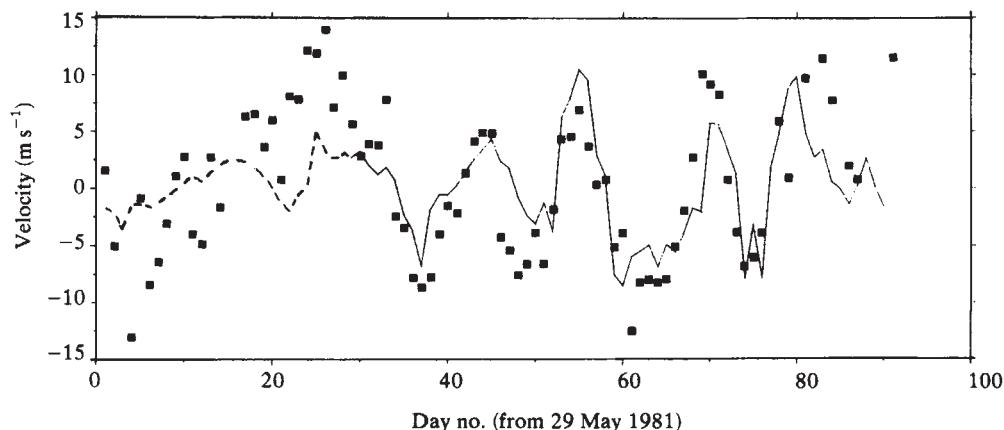
Claverie *et al.*¹ point out that their data from previous years confirm the presence of the 13-day velocity signal. However, the observed period ranges from 13.08 days in 1977 to 16.53 days in 1980 and lack of knowledge has prevented comparisons similar to that presented above for these data.

In our calculations several numerical parameters are introduced, the uncertainty of which will affect the calculated velocity signal. It seems likely that the largest inaccuracies are introduced by the uncertainty in the measured sunspot area and in the line intensity within the penumbra. The uncertainty in sunspot area will tend to decrease the correlation between the observed and the calculated values. If any large scale velocity field affects the measurements it seems likely that this field coincides in heliographic longitudes with the sunspot groups. We have carried out several test such as including the limb effect⁹ and changing the line broadening, but no large deviations from the present model were found.

Based on our simple model calculation we note that variations in the apparent velocity signal will result naturally from the passage of sunspots over the solar disk. It appears that the model is able to account for the observed velocity signal¹ during a time span of 88 days. We conclude that further studies are required, for instance by a two-dimensional spectrometer¹⁰, before any firm conclusion regarding the solar core rotation can be made.

The paper by Durrant and Schröter¹¹ was brought to our attention since submission of this letter. Their conclusion that the presence of active regions will modulate the observed integrated velocity signal is consistent with ours. However,

Fig. 2 The calculated velocity signal as caused by the passage of sunspots over the solar disk is plotted against time and may be compared with the Izana observations¹. Note that a reasonable fit to the observations is also obtained during the first 30 days (dashed line) where the suggested 13-day period does not fit the data.



contrary to our findings they argue qualitatively that sunspots cover too small an area to modulate significantly the velocity signal. Although our calculations show that the influence of sunspots alone may account for the observations, possible additional modulations of the integrated velocity signal by other parts of the active region cannot be excluded.

Received 14 December 1982; accepted 15 February 1983.

1. Claverie, A. *et al.* *Nature* **299**, 704–707 (1982).
2. Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Mon. Not. R. astr. Soc.* **185**, 19–22 (1978).
3. Pierce, A. K., Slaughter, C. D. & Weinberger, D. *Solar Phys.* **52**, 179–189 (1977).
4. Albrechtsen, F. & Maltby, P. *The Physics of Sunspots* (eds Cram, L. & Thomas, J. H.) 127–139 (Sacramento Peak Observatory, 1981).
5. Maltby, P. *Solar Phys.* **26**, 76–82 (1972).
6. Scherrer, P. H., Wilcox, J. M. & Svalgaard, L. *Astrophys. J.* **241**, 811–819 (1980).
7. Snider, J. L. *Phys. Rev. Lett.* **28**, 853–855 (1972).
8. Beckers, J. M. *Astrophys. J.* **213**, 900–905 (1977).
9. Howard, R., Boyden, J. E. & LaBonte, B. J. *Solar Phys.* **66**, 167–185 (1980).
10. Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Solar Phys.* **74**, 503–508 (1981).
11. Durrant, C. J. & Schröter, E. H. *Nature* **301**, 589–591 (1983).

Solar atmospheric temperature inhomogeneities induce a 13-day oscillation in full-disk Doppler measurements

M. G. Edmunds

Department of Applied Mathematics and Astronomy, University College, PO Box 78, Cardiff CF1 1XL, UK

D. O. Gough

Institute of Astronomy, and Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Claverie *et al.*¹ have reported a 13-day oscillation in full-disk solar spectrum-line shifts, which they claim to be evidence for rapid rotation of the Sun's core. We point out here that the passage of active regions across the disk could be responsible for oscillations of this kind. Variations in line shifts and in irradiance are expected with frequencies comparable to both the photospheric rotation rate and to higher harmonics. For line-shift oscillations the power in the second harmonic is substantially greater than it is for irradiance variations. Therefore a prominent 13-day component is not surprising, especially as we know that the second harmonic is strong in the sunspot distribution². We estimate the amplitude, period and phase that would be expected to have been found in line-shift data, obtaining results comparable with those of Claverie *et al.*¹. We suggest that the 13-day oscillation is a direct consequence of the inhomogeneity of the solar surface. Our analysis complements and extends a recent study by Durrant and Schröter³.

Because the surface of the Sun is rotating, any east–west asymmetry in the intensity of the radiation produces a shift in a spectrum line in light integrated over the entire disk. We illustrate the phenomenon using a simple model of the solar atmosphere. We assume that the axis of rotation is perpendicular to the line of sight, and specify any point of the surface by the polar angles (θ, ϕ) about that axis, orientated such that $(\pi/2, 0)$ is at disk centre, and we presume that the only significant line shift is a Doppler shift resulting from the synodic surface angular velocity $\Omega(\theta)$. We now consider the surface of the Sun to be partially covered by patches such as sunspots or plages⁴, whose only function is to modify the intensity $I(\theta, \phi; \lambda; t)$ at wavelength λ and time t . Such a representation is supported by Willson's⁵ analysis of SMM-ACRIM solar irradiance data. We ignore systematic velocity variations in active regions, because their influence on the full-disk Doppler measurements is substantially smaller than the principal effect we are considering. The patches are distributed such that $A(\theta, \phi; t)$ is the proportion of any small area centred at (θ, ϕ) that is covered by them. The operator Δ signifies the excess of a quantity in the patches over the quiet Sun, and an overbar, the average over the visible disk.

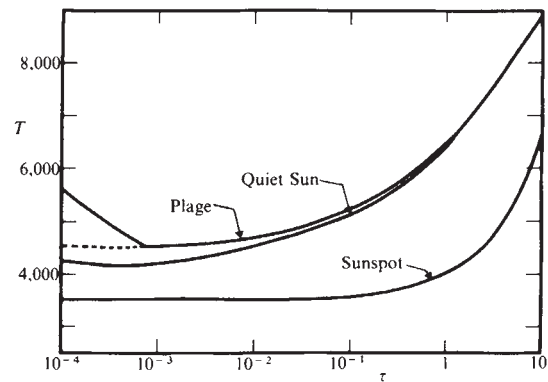


Fig. 1 The solid lines represent the T - τ relations used to model the quiet Sun, sunspots and plages. The broken line represents an alternative plage model, which predicts contributions to V that are $\sim 10\%$ lower than those quoted in the text.

The quantity used by Claverie *et al.*¹ to measure velocity is essentially⁶

$$\mathcal{R}(t) = \frac{\bar{I}(\lambda_-) - \bar{I}(\lambda_+)}{\bar{I}(\lambda_-) + \bar{I}(\lambda_+)} \quad (1)$$

where $\lambda_{\pm} = \lambda_0 \pm \lambda_1$; $\lambda_0 = 769.9$ nm is the wavelength at line centre in the frame of the detector, and $\lambda_1 = 6.65$ pm, corresponding to a Doppler velocity of 2.69 km s⁻¹, and is such that $\lambda_{\pm}$ are near the wavelengths where the slope of $\bar{I}(\lambda)$ is greatest (we omit the argument t from $\bar{I}$). As the velocity of the detector relative to the Sun varies with time, so does the response $\mathcal{R}$ to a small solar line shift $\delta\lambda$. However, Claverie *et al.*¹ have made some correction for this. Therefore, in order to predict their corrected data, we regard λ_0 as the centre of the solar line in the absence of rotation, and ignore second and higher derivatives of $\bar{I}(\lambda)$ at $\lambda = \lambda_{\pm}$. In that case the response $\mathcal{R}$ may easily be converted to a velocity V away from the observer using the Doppler formula on the integrated line profile:

$$V \approx \frac{1}{2}c[\bar{I}(\lambda_+) + \bar{I}(\lambda_-)]\lambda_0^{-1}\bar{S}^{-1}\mathcal{R} \quad (2)$$

where $S(\theta, \phi; t) = \frac{1}{2}[I'(\lambda_+) - I'(\lambda_-)]$ where c is the velocity of light and the prime denotes differentiation with respect to λ . A shift $\delta\lambda(\theta, \phi) = c^{-1}\lambda_0 R \sin \theta \sin \phi \Omega(\theta)$ of the line, where R is radius of the Sun, therefore yields

$$V \approx R\bar{S}^{-1}\bar{S}\Omega \sin \theta \sin \phi \quad (3)$$

If I were a symmetric function of ϕ , then $V = 0$. However, an asymmetrical distribution of patches produces an integrated velocity signal

$$V = (\pi\bar{S})^{-1}R \int_0^\pi \Omega(\theta) \sin^3 \theta d\theta \int_{-\pi/2}^{\pi/2} A \Delta S \sin \phi \cos \phi d\phi \quad (4)$$

which in general is non-zero. It also produces a variation δF in the solar constant F :

$$\frac{\delta F}{F} = (\pi\bar{\mathcal{J}})^{-1} \int_0^\pi \sin^2 \theta d\theta \int_{-\pi/2}^{\pi/2} A \Delta \mathcal{J} \cos \phi d\phi \quad (5)$$

where $\mathcal{J}(\theta, \phi; t)$ is the white-light intensity.

At each co-latitude we now develop the patch distribution in a Fourier series in the lagrangian azimuthal angle $\tilde{\phi} = \phi - \Omega t$:

$$A(\theta, \phi; t) = \sum_{n=0}^{\infty} A_n(\theta, t) \cos(n\tilde{\phi} + \epsilon_n) \quad (6)$$

We assume that in a frame rotating locally with the solar surface the patch distribution changes slowly, so that the time dependence of A_n can be ignored; this seems to be a good first approximation for $n \geq 2$. Note also that during the time when Claverie *et al.*¹ made their observations, the sunspots were concentrated in two narrow bands of latitude, with a mean magnitude of $0.074\pi \equiv \pi/2 - \Theta$ and a standard deviation of only 0.033π . We therefore ignore the finite widths of the bands, and set

$$A_n(\theta, t) + A_n(\pi - \theta, t) = a_n \delta(\theta - \Theta), \quad 0 \leq \theta \leq \pi/2 \quad (7)$$

where $\delta(\theta)$ is the Dirac delta function and the a_n are constants. The velocity signal and solar constant variations can then be evaluated easily in terms of the Fourier coefficients. Assuming limb-darkening variations of the form

$$\begin{aligned}\Delta S &= (1 - u_s + u_s \cos \psi) \Delta S_0 \\ \Delta \mathcal{F} &= (1 - u_i + u_i \cos \psi) \Delta \mathcal{F}_0\end{aligned}\quad (8)$$

where $u_s, u_i, \Delta S_0, \Delta \mathcal{F}_0$ are constants and ψ is the angle subtended from disk centre, the results are

$$\begin{aligned}V &= \frac{1}{2} R \Omega_0 (\pi \bar{S})^{-1} \Delta S_0 \sin^3 \Theta \\ &\times \sum_{n=0}^{\infty} [(1 - u_s) \alpha_{n,2}^- + \frac{1}{2} u_s (\alpha_{n,1}^- + \alpha_{n,3}^-) \sin \Theta] \\ &\times a_n \sin(n \Omega_0 t - \varepsilon_n)\end{aligned}\quad (9)$$

$$\begin{aligned}\frac{\delta F}{F} &= (\pi \bar{\mathcal{F}})^{-1} \Delta \mathcal{F}_0 \sin^2 \Theta \\ &\times \sum_{n=0}^{\infty} [(1 - u_i) \alpha_{n,1}^+ + \frac{1}{2} u_i (\alpha_{n,0}^+ + \alpha_{n,2}^+) \sin \Theta] \\ &\times a_n \cos(n \Omega_0 t - \varepsilon_n)\end{aligned}\quad (10)$$

where $\Omega_0 = \Omega(\Theta)$ and

$$\alpha_{n,m}^{\pm} = \begin{cases} \frac{\pi}{2} \pm \frac{\pi}{2} & \text{if } n = 0 \text{ and } m = 0 \\ \left[\frac{1}{n-m} \pm \frac{(-1)^m}{n+m} \right] \sin \frac{1}{2}(n-m)\pi & \text{if } n^2 - m^2 \neq 0 \\ \frac{\pi}{2} & \text{if } n = m \neq 0 \end{cases} \quad (11)$$

Two simple properties can be observed from these formulae. First, for the n th Fourier component the maximum in V occurs $\pi/(2n\Omega_0)$ after or before maximum δF , according to whether $\Delta S_0/\Delta \mathcal{F}_0$ is positive or negative: this results from the fact that, as is apparent from symmetry considerations, the influence of a surface inhomogeneity on V is an odd function of ϕ , whereas the influence on δF is even. Second, the ratio r_V of the amplitudes of the second and first harmonics of V exceeds the similar ratio r_F for δF . For example, if limb darkening is ignored, $r_V: r_F = 9\pi^2/32$, which would imply that compared with the fundamental the power in the second harmonic is 8 times greater for V than it is for δF .

To estimate $\bar{S}^{-1} \Delta S_0$ and $\bar{\mathcal{F}}^{-1} \Delta \mathcal{F}_0$ we constructed three model atmospheres in hydrostatic equilibrium (with no turbulent pressure) and local thermodynamic equilibrium (LTE) to represent the quiet Sun, the centre of a sunspot umbra, and a plage. The atmospheres were computed from given $T-\tau$ relations (where T is temperature and τ is continuum optical depth at $\lambda = 500$ nm) for $\tau \geq 10^{-4}$, using the approximations to the continuum opacity given by Gingerich⁷. Chemical abundances were taken from refs 8–11. We then computed profiles of the potassium 769.9 nm line, both at disk centre and in light integrated over the disk, so that centre-to-limb variations could be estimated. Here, as the height dependence of the microturbulent broadening parameter is still not well determined in active regions, sunspots or the quiet Sun, we adopted for simplicity a height-independent microturbulent velocity of 0.5 km s^{-1} throughout. We took pressure damping to be twice that given by the van der Waals formula, and determined a gf factor of 0.69 by requiring that the equivalent width of the line at disk centre in the model of the quiet Sun be 15 pm, in agreement with observation. For the quiet Sun we adopted the $T-\tau$ relation of Model C of Vernazza *et al.*¹², for sunspots we used the model of Albregtsen and Maltby¹³ extended to $\tau = 10^{-4}$ with the $T-\tau$ relation of Avrett¹⁴, and for the plage we used a relation estimated from Model P1 of Lemaire *et al.*¹⁵ (see Fig. 1). We also considered a plage model estimated as a compromise between Model F of Vernazza *et al.*¹² and Model P1 of Lemaire *et al.*¹⁵ (broken line in Fig. 1).

We assumed the plages to be horizontally uniform. A sunspot, however, varies substantially from the centre of the umbra to the edge of the penumbra. Nevertheless, it is evident from the analysis of Flå *et al.*¹⁶ that a linear interpolation of the properties from umbral to quiet conditions gives a fair approximation to the intermediate properties. Therefore, we simply diluted the umbral model by multiplying the values of ΔS_0 and $\Delta \mathcal{F}_0$ by the factor $0.32/0.82 = 0.39$, which is the ratio of the mean intensity deficit in a typical sunspot¹⁷ to the value computed at the centre of the umbra. Thus for the average over sunspots we obtained $\bar{S}^{-1} \Delta S_0 = -0.47$ with $u_s = 0.55$, and $\bar{\mathcal{F}}^{-1} \Delta \mathcal{F}_0 = -0.39$. The corresponding values for plages are -0.13 , -0.54 and 0.01 , respectively. We adopted a white-light limb-darkening parameter $u_i = 0.5$ for both sunspots and plages.

Note that at $\lambda_{\pm}$ the slope of the potassium line is less steep in both sunspots and plages. Consequently $\bar{S}^{-1} \Delta S_0$ has the same sign in both cases, and the two contributions to V reinforce each other. The value of $\bar{S}^{-1} \Delta S_0$ for the cooler plage model is $\sim 10\%$ less than that quoted above.

We have estimated V separately from variations in the solar constant measured by SMM-ACRIM⁴ and from the sunspot distribution. The latter were provided by Hoyt in the form of a predicted deficit in the solar flux produced by sunspot blocking. Both sets of data were analysed for the interval 28 June–22 August 1981, which is the interval considered the most thoroughly by Claverie *et al.*¹. From Fourier transforms of the data we estimated the coefficients a_n using equation (10). Hudson and Willson¹⁸ have inferred from the SMM-ACRIM data that the actual variation of the solar constant is about 70% of that produced by sunspot blocking alone; we have adopted this factor in our computations. The contribution to V from sunspots can then be obtained directly from equation (9). Plages are found in close association with sunspots, and we assume that the plage distribution is similar to that of the sunspots. Thus for plages we multiplied the a_n by a constant factor β , and adopted $\beta = 7$, a value characteristic of the ratio of the total areas occupied by plages and sunspots at times of high solar activity¹⁷, particularly during the time interval considered here (D. Hoyt, personal communication).

The power spectrum of the sunspot data has peaks which, aside from the first, are approximately uniformly spaced. This is consistent with what one would expect from a pattern which we know changes on a time scale of roughly two rotation periods. The power in the first four harmonics of the spectrum of V inferred from equation (9) is in the ratio $0.67:1:0.26:0.05$. Thus V is dominated by the second harmonic, which has a period of 13.8 days and an amplitude of 3.7 m s^{-1} (of which 1.0 m s^{-1} came from sunspots and 2.7 m s^{-1} from plages). Velocity minimum, for this component, occurred on 12 August. From the SMM-ACRIM data we obtained a period of 13.1 days and an amplitude of 4.0 m s^{-1} , with velocity minimum on 11 August. Claverie *et al.*¹ analysed separately their data from Izaña and Haleakala, the latter being apparently the more noisy. From the two months of independent data they obtained periods of 13.15 and 13.38 (± 0.2) days, each with associated velocity minima on 13 August, and amplitudes of 6.6 m s^{-1} and 9.8 m s^{-1} , respectively.

The periods and phases of the 13-day component of the full-disk spectrum-line shift induced by plages and sunspots that we have computed are in quite good agreement with the observations of Claverie *et al.*¹. Durrant and Schröter³ have drawn a similar conclusion from a detailed analysis of the calcium index, which is probably a more reliable measure of the brightness-weighted plage area than ours. However, accordance of period and phase alone is insufficient to demonstrate that active regions are directly responsible for the line-shift oscillations, because both might be symptoms of something else. It is necessary also to know the amplitude of the line shift that the intensity inhomogeneities induce. Our estimate of that amplitude shows it to be of the same order of magnitude as that observed.

Much of the discrepancy between our results and observation probably arises from our misrepresentation of the structure and distribution of plages. There is considerable uncertainty in the plage models¹⁵. Moreover, our simple rule that the plage area is β times the sunspot area may be unreliable. We suspect that this, coupled with our neglect of the evolution of the sunspots and plages, is responsible for the fact that our second harmonic is not as prominent as that observed. We have not pursued this matter because Durrant and Schröter³ have produced a power spectrum very similar to that of Claverie *et al.*¹ from spatially resolved data. One should appreciate, however, that the modification to the 769.9 nm spectrum line may not be a linear function of the visual estimate of the plage brightness that Durrant and Schröter used, and that therefore the relative amplitudes they inferred may not be correct.

Our neglect of non-LTE effects high in the atmosphere might have severely impaired our results. However, deviations from LTE influence mainly the line core¹⁹; the shape of the line near $\lambda = \lambda_{\pm}$ is formed deeper in the atmosphere, where deviations from LTE appear to be small. This is verified by the quite minor difference between the predictions of the two plage models illustrated in Fig. 1. More serious, however, may be our assumption that microturbulent broadening is the same in sunspots and plages as it is in the quiet Sun. The amplitude of V could possibly be engineered to agree more closely with observation by adjusting this device within plausible limits.

We conclude that a substantial fraction of the 13-day line-shift variation reported by Claverie *et al.*¹ is a direct result of atmospheric temperature inhomogeneities giving non-uniform weight to the Doppler shift from the rotation of the surface of the Sun. Our analysis does not preclude a further contribution from other possible large-scale components of the surface velocity²⁰. Horizontal velocities associated with giant cells, for example, may also be important. One would expect flow structure to be correlated with the pattern of magnetic field, and hence with active regions; that could reinforce the 13-day full-disk signal.

Finally, we have carried out a similar analysis of the SMM-ACRIM and sunspot data in the interval 1 July to 31 August 1982. Once again the second harmonic dominates; from the SMM data we deduce a period of 13.9 days with an implied velocity minimum on 13 August, and from the sunspot data we deduce a period of 14.3 days with velocity minimum on 15 August, and the same amplitude as in 1981.

We thank Dr D. Hoyt for supplying sunspot data from the Space Environment Laboratory (NOAA, Boulder, USA), Dr R. C. Willson for unpublished SMM-ACRIM data, G. Athay, E. H. Avrett, J. A. Eddy, J. W. Harvey, P. H. Scherrer, N. O. Weiss and J. M. Wilcox for interesting discussions, H. B. van der Raay for pointing out an error in our original computation of the 1982 sunspot phase, and Dr C. J. Durrant for suggesting improvements to the original typescript.

Received 10 January; accepted 3 March 1983.

- Claverie, A. *et al.* *Nature* **299**, 704–707 (1982).
- Knight, H. W., Schatten, K. H. & Sturrock, P. A. *Astrophys. J.* **327**, 153–156 (1979).
- Durrant, C. J. & Schröter, E. H. *Nature* **301**, 589–591 (1983).
- Willson, R. C. *et al.* *Science* **211**, 700–702 (1981).
- Willson, R. C. *J. geophys. Res.* **87**, 4319–4326 (1982).
- Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Mon. Not. R. astr. Soc.* **185**, 1–17 (1978).
- Gingerich, O. *Proc. 2nd Harvard-Smithsonian Conf. on Stellar Atmospheres* (SAO Spec. Rep. No. 167, 1964).
- Warner, B. *Mon. Not. R. astr. Soc.* **138**, 229–243 (1968).
- Foy, R., *Astr. Astrophys.* **18**, 26–38 (1972).
- Lambert, D. L. *Mon. Not. R. astr. Soc.* **182**, 249–272 (1978).
- Lambert, D. L. & Luck, R. E. *Mon. Not. R. astr. Soc.* **183**, 79–100 (1978).
- Vernazza, J. E., Avrett, E. H. & Loeser, R. *Astrophys. J. Suppl.* **45**, 635–725 (1981).
- Albregtsen, F. & Maltby, P. in *The Physics of Sunspots* (eds Cram, L. E. & Thomas, J. H.) 127–139 (Sacramento Peak Observatory, 1981).
- Avrett, E. H. *The Physics of Sunspots* (eds Cram, L. E. & Thomas, J. H.) 235–255 (Sacramento Peak Observatory, 1981).
- Lemaire, P. *et al.* *Astr. Astrophys.* **103**, 160–176 (1981).
- Flå, T., Osherovich, V. A. & Skumanich, A. *Astrophys. J.* **261**, 700–709 (1982).
- Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
- Hudson, H. S. & Willson, R. C. in *The Physics of Sunspots* (eds Cram, L. E. & Thomas, J. H.) 434–445 (Sacramento Peak Observatory, 1981).
- de la Reza, R. & Müller, E. A. *Sol. Phys.* **43**, 15–32 (1975).
- Dicke, R. H. *Nature* (in the press).

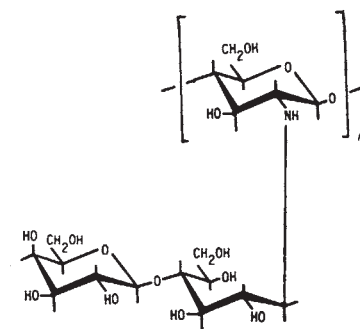
Unusual rheology of a branched, water-soluble chitosan derivative

Mansur Yalpani*†, Laurance D. Hall‡, Marvin A. Tung§ & Donald E. Brooks*‡

Departments of Pathology*, Chemistry‡, and Food Sciences§, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

The rheological properties of polysaccharides are exploited for many industrial applications which have created an increasing demand for new inexpensive derivatives with improved performance¹. *N*-Akylation of the linear polysaccharide chitosan with various aldehydo-, keto- or lactone sugars affords branched-chain, water-soluble derivatives of high molecular weight for which the extent of branching as well as the branch conformation, length and type can readily be modified by appropriate choice of reagents and reaction conditions². The products exhibit a correspondingly broad spectrum of rheological properties. We report here on some unusual non-newtonian features of one of these derivatives, 1-deoxylactit-1-yl chitosan, which displays low-shear newtonian behaviour, a medium-shear viscosity increase (dilatancy), and a high-shear viscosity drop (pseudoplasticity). In conditions of steady shear the shear stress-time function varies in the different shear rate regimes. A 2% solution of this derivative exhibits unique oscillatory behaviour under steady shear at the transition point between the dilatant and pseudoplastic regimes.

1-Deoxylactit-1-yl chitosan, *1*, a comb-like polymer, is obtained in high yields and with high degrees of substitution (d.s., essentially quantitative) by reductive alkylation of the



2-deoxy-2-amino functions of chitosan, using lactose and sodium cyanoborohydride². Aqueous solutions of this derivative reveal a complex viscosity profile as illustrated in Fig. 1. A 1% solution produces newtonian characteristics at low shear rates ($\dot{\gamma} \leq 10 \text{ s}^{-1}$). At moderate shear rates ($10 \text{ s}^{-1} < \dot{\gamma} < 40 \text{ s}^{-1}$), derivative *1* is dilatant with an approximately fivefold viscosity increment. This dilatant regime gives way to a pseudoplastic one at higher shear rates ($\dot{\gamma} > 40 \text{ s}^{-1}$).

From Fig. 1 it is also evident that the shear rate ($\dot{\gamma}_{\text{max}}$) at which the viscosity maximum (η_{max}) is attained shifts to progressively lower values with increasing polymer concentration. Conversely, decreasing concentrations lead to attenuation of the dilatant features, until at 0.5% essentially newtonian behaviour is observed.

Interestingly, pronounced oscillations in shear stress are found for the 2% solution, as indicated in Figs 1 and 2A, when the shear rate is increased from $\dot{\gamma}_{\text{max}}$ to the next higher, critical value, $\dot{\gamma}_{\text{Cl}}$. At this value ($\dot{\gamma}_{\text{Cl}} = 4.75 \text{ s}^{-1}$; 25°C), the apparent viscosity oscillates with a constant frequency of $\sim 0.296 \pm 0.033 \text{ Hz}$ without any appreciable damping being apparent over the periods (few minutes) we have investigated. The

† Present address: B. C. Research, Chemical Technology Division, 3650 Wesbrook Mall, Vancouver, British Columbia, Canada V6S 2L2.

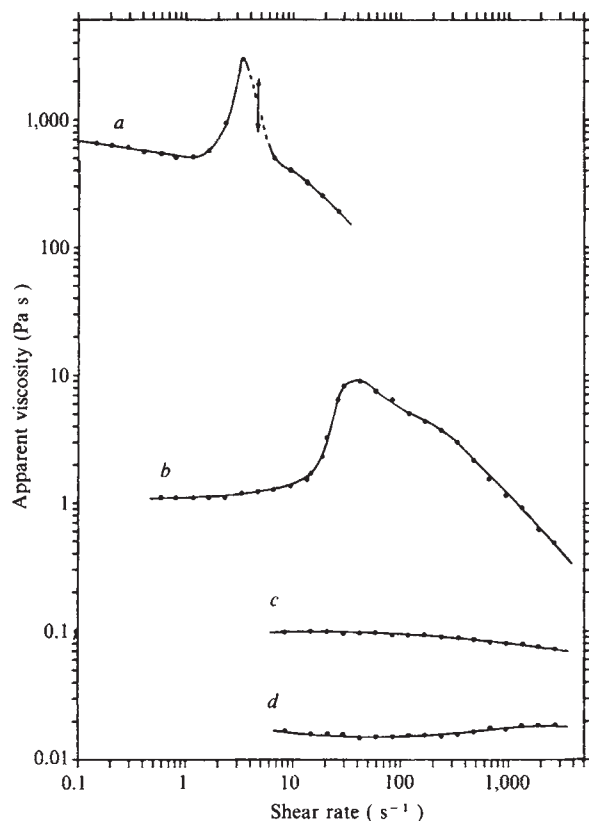


Fig. 1 Apparent viscosity (shear stress/shear rate) versus shear rate curves for aqueous solutions of 1-deoxylactit-1-yl chitosan at various concentrations (wt/wt): *a*, 2%; *b*, 1%; *c*, 0.5%; *d*, 0.1% at 25.0 °C. The bar in curve *a* indicates the oscillation amplitude at a shear rate of 4.75 s^{-1} (see also Fig. 2A). The polymer was prepared according to previously reported methods². The degree of substitution was determined by elemental microanalysis to be 0.97 ± 0.01 with 0.03 ± 0.01 residual *N*-acetyl substitution; any polyelectrolyte characteristics should therefore be confined to $\leq 3\%$ of the monomeric residues. The number average molecular weight (M_n) derived from sodium boro [^3H] hydride end-group analysis was 1.25×10^6 ; this value was in agreement with the expected increase in molecular weight based on a M_n value of 4.17×10^5 for the native polymer (the latter being consistent with reported values^{14,15}). Aqueous solutions of the polymers were prepared at ambient temperature by dilution of a stock and contained 0.02% sodium azide. Viscosity determinations were performed on a Brabender Rheotron viscometer using the A_1 measuring system (coaxial cylinders with rotating outer cylinder; dimensions: length 8.00 cm, gap 1.0 mm, $r_o = 2.80$ cm, $r_i = 2.70$ cm) with springs C (curve *a*), A and B (curve *b*), and B (curves *c*, *d*). The natural frequencies of the A_1 fixture were (spring): 4.57 Hz (A), 10.21 Hz (B), and 32.28 Hz (C). Shear rate values were derived from instrument settings and have not been corrected for non-newtonian flow. Shear stress values used in calculating apparent viscosities were obtained in all cases after attainment of steady values, and were reproducible (within a few percent) in successive cycles of increasing or decreasing shear rate, as well as after brief periods (0.5–2.0 min) without application of shear. Similarly, only minor differences ($\sim 10\%$) in viscosity levels were detectable when fresh solutions were compared with those on which a complete set of measurements had been made as a function of shear rate, or those which had been stored in the presence of bacteriostatic agents for periods of up to seven weeks. The observed phenomena did not appear to be dependent on the preparation.

peak-to-peak amplitude of the oscillation ranges from a value which corresponds approximately to that at the onset of the pseudoplastic region to one which is smaller than η_{max} . This phenomenon, which is not observed for the 1% solution, presumably reflects a transition between the two major structural domains of the polymer in solution which produce dilatant and pseudoplastic flow, respectively. Further support for this supposition derives from the fact that another, albeit much less

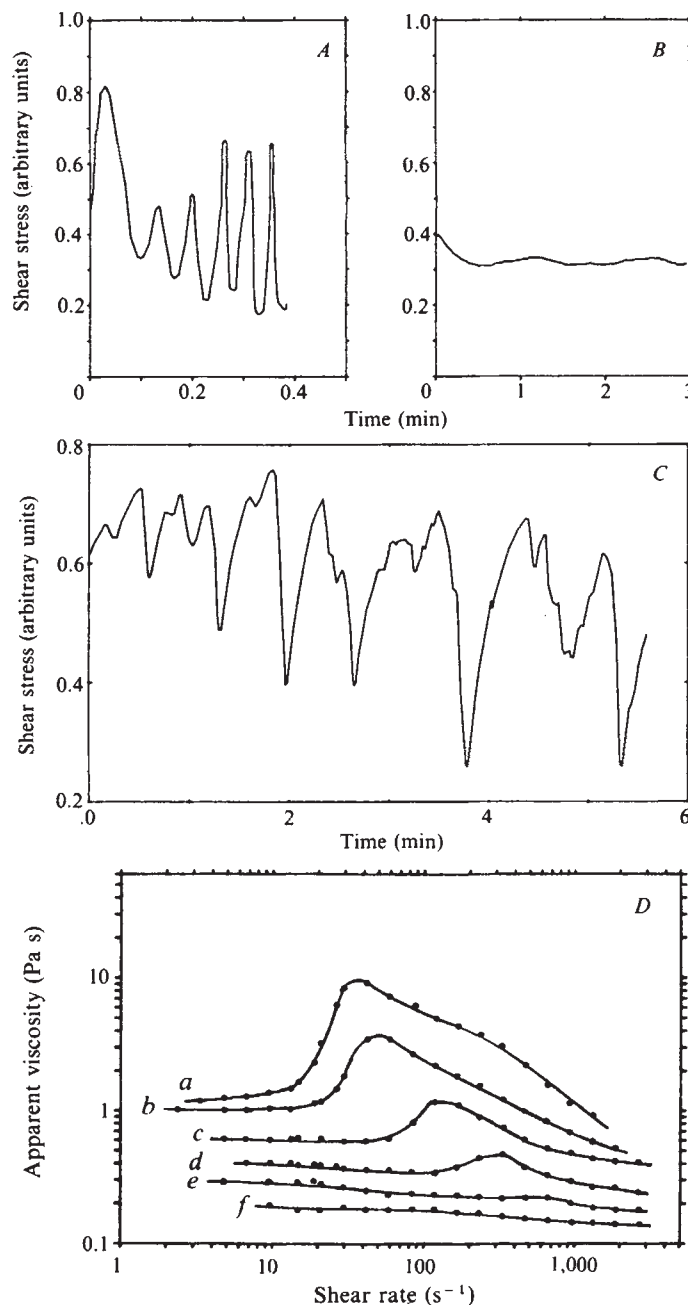


Fig. 2 Ambient temperature (25.0 °C) shear stress-time dependencies at steady shear (polymer concentration, scale factor): A, 4.75 s^{-1} (2%, $10\times$); B, 2.37 s^{-1} (2%, $3\times$); C, 0.30 s^{-1} (0.91%, $1\times$) in the presence of sodium chloride (9.2% wt/wt). D, Effect of temperature on apparent viscosity-shear rate profiles of 1% solution at *a*, 25.0 °C; *b*, 27.5 °C; *c*, 30.0 °C; *d*, 35.0 °C; *e*, 43 °C; *f*, 50.0 °C. Traces in A–C represent direct instrument readouts and are proportional to shear stress (arbitrary units; fraction of full scale in C). Oscillations in B are significantly larger than instrumental noise; zero time was chosen arbitrarily in C. Note that addition of salt in C was accompanied by an approximately twofold increase in solution viscosity.

distinct, oscillation is observed at the onset of the dilatant regime. The peak-to-peak oscillation at this critical shear rate ($\dot{\gamma}_{c2} = 2.37 \text{ s}^{-1}$; 25 °C) amounts only to $\sim 10\%$ of the viscosity minima values (Fig. 2B). Whereas in the previous transition the oscillation frequency and the critical shear rate correspond in magnitude, the oscillation frequency in this case ($\sim 0.012 \text{ Hz}$) exceeds $\dot{\gamma}_{c2}$ by more than an order of magnitude. A particularly unusual oscillatory pattern was observed for a 0.9% solution of the polymer on addition of 9.2% sodium chloride. As shown in Fig. 2C, a rather irregular oscillation with substantial variations in peak-to-peak amplitude was obtained at a shear rate

value of 0.30 s^{-1} . The fact that the natural frequencies of the measuring system (see legend Fig. 1) are relatively high would preclude any artefacts due to fixture oscillation. It also seems unlikely that the oscillation phenomena can be ascribed to polyelectrolyte effects, since the addition of salt did not result in a lowering but in an increase (see legend Fig. 2) of the apparent viscosity.

Another noteworthy feature of 1-deoxylactit-1-y1 chitosan is the variety of viscosity-time dependencies (at constant shear) it exhibits in the different shear rate regimes. Thus, for the 1% solution, the shear stress is virtually time-independent in the low-shear newtonian regime, the shear stress decays to a plateau with time in the dilatant regime, and below 470 s^{-1} , where pseudoplasticity is observed, the shear stress increases with time to a constant value. Above 470 s^{-1} , the sequence of increasing, constant, or decaying stress preceding attainment of a steady value was again found with progressively increasing shear rates. Similar trends are displayed by 2% solutions, but not by lower concentrations ($\leq 0.5\%$) of the polymer derivative.

Figure 2D shows the effects of temperature on the viscosity of a 1% solution of the branched chitosan derivative. Not surprisingly, the $\dot{\gamma}_{\text{max}}$ value increases with temperature and, concurrently, the dilatant feature disappears gradually, until, at $\sim 43^\circ\text{C}$, essentially newtonian flow is again observed. Tem-

perature increments also result in corresponding shifts of the various viscosity-time patterns to higher shear rates.

We believe that the complex rheological properties of 1-deoxylactit-1-y1 chitosan are unique for a homogeneous aqueous polysaccharide solution. In particular, we are not aware of reports on steady shear viscosity oscillations in any other type of system. The native chitosan (in dilute aqueous acetic acid)^{3,4}, guaran⁵, xanthan gum⁶, and indeed most other industrially important polysaccharides⁷ exhibit monotonic pseudoplasticity with increasing shear. While simple dilatancy has been reported for aqueous dispersions of starch and other gums⁸, viscosity maxima and other, similarly complex features have so far only been observed for some inorganic and organic polymers, including boron-siloxanes⁹, polyisobutylene¹⁰, and polystyrene¹¹⁻¹³, in organic solutions.

The complex viscosity profiles and oscillation patterns observed here may be indicative of the formation and disintegration of different types of solution structures. In the absence of further independent evidence, however, proposals of more specific models for the structures involved or for the mechanisms which trigger the transitions between them would seem unjustified.

This work was supported by grants from the MRC of Canada. We thank Dr K. Pinder for helpful discussions.

7. Morris, E. R. & Ross-Murphy, S. B. in *Techniques in Carbohydrate Metabolism*, 1-46 (Elsevier, Amsterdam, 1981).

8. Eastwood, A. R. & Barnes, H. A. *Rheol. Acta* **14**, 795-800 (1975).

9. Trapeznikov, A. A., Frolova, Y. A. & Zatspeina, T. I. *Polym. Sci. USSR* **17**, 1948-1954 (1975).

10. Trapeznikov, A. A. & Pylayeva, A. T. *Polym. Sci. USSR* **17**, 2013-2022 (1975).

11. Peterlin, A. & Turner, D. T. *J. chem. Phys.*, **38**, 2315-2316 (1963).

12. Layec-Raphalen, M. N., & Wolff, C. J. *Non-Newtonian Fluid Mech.* **1**, 159-173 (1976).

13. Cheng, D. C. H. *Nature* **245**, 93-95 (1973).

14. Muzzarelli, R. A. A. *Natural Chelating Polymers* (Pergamon, Oxford, 1973).

15. Wu, A. C. M., Bough, W. A., Conrad, E. C. & Alden, K. E. *J. Chromatogr.* **128**, 87-99 (1976).

1. Whistler, R. L. (ed.) *Industrial Gums* 2nd edn (Academic, New York, 1973).

2. Hall, L. D. & Yalpani, M. *JCS chem. Commun.*, 1153-1154 (1980).

3. Sklyar, A. M. *et al. Polym. Sci. USSR* **23**, 1396-1403 (1981).

4. Lang, E. R., Kienle-Sterzer, C. A., Rodriguez-Sanchez, D. & Rha, C., *Proc. 2nd int. Conf. on Chitin and Chitosan* (eds Hirano, S. & Tokura, S.) 34-38 (The Japanese Society of Chitin and Chitosan, Tottori, 1982).

5. Sharman, W. R., Richards, E. L. & Malcolm, G. N. *Biopolymers* **17**, 2817-2833 (1978).

6. Chen, C. S. H. & Sheppard, E. W. *Polym. Engng Sci.* **20**, 512-516 (1980).

Combined Sr-O isotope relationships and petrogenesis of Andean volcanics of South America

Bernard Dérulle*, Russell S. Harmon† & Stephen Moorbath‡

* Department of Earth Sciences, Faculty of Science, University of Yaounde, BP 812, Yaounde, Cameroun

† Department of Geological Sciences, Southern Methodist University, Dallas, Texas 75275, USA

‡ Department of Geology and Mineralogy, Oxford University, Parks Road, Oxford OX1 3PR, UK

Active andesitic-type volcanism in the southern Andes (lat. $42^\circ\text{--}36^\circ\text{S}$) is characterized by low Sr- and O-isotopic compositions which fall in the range $^{87}\text{Sr}/^{86}\text{Sr} = 0.7036\text{--}0.7043$, $\delta^{18}\text{O} = 5.2\text{--}6.8\%$. In contrast, volcanic centres in the Central Andes (lat. $24\text{--}21^\circ\text{S}$) are significantly enriched in both ^{87}Sr and ^{18}O . We present the first combined Sr- and O-isotope study of the southern volcanic zone (SVZ) in Chile (crustal thickness $<40 \text{ km}$)¹ and of the southern part of the central volcanic zone (CVZ) in the borderlands of Bolivia, Chile and Argentina (crustal thickness $\sim 60\text{--}70 \text{ km}$)¹. Only one of 13 volcanic centres studied (Lascar) exhibits a correlated increase in both ^{87}Sr and ^{18}O parallel with the basalt-andesite-rhyolite compositional change, indicative of contamination by upper crust through a fractional crystallization-assimilation mechanism. The other centres of the CVZ, for which there is an overall Sr-O correlation, document the variable contamination of mantle-derived magmas at depth in the lower crust which had been thickened by Mesozoic and earlier plutonism.

Active volcanism in the Andes occurs in three linear zones parallel to the western coast of South America: a northern zone (NVZ, $5^\circ\text{N--}2^\circ\text{S}$), a central zone (CVZ, $16\text{--}28^\circ\text{S}$), and a southern zone (SVZ, $33\text{--}52^\circ\text{S}$), each with calc-alkaline associations of distinctive chemical and isotopic compositions. Published $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{18}\text{O}/^{16}\text{O}$ measurements on the NVZ of

Ecuador mostly yield low values ($^{87}\text{Sr}/^{86}\text{Sr} = 0.7036\text{--}0.7043$, $\delta^{18}\text{O} = 6.5\text{--}7.7\%$), whilst the CVZ is characterized by ^{87}Sr - and ^{18}O -enrichment which clearly demonstrates the occurrence of complex magma-crust interaction processes (for south Peru, $0.7054\text{--}0.7076$ and $7.0\text{--}8.6\%$; for north Chile, $0.7056\text{--}0.7070$ and $8\text{--}10.8\%$; for north-west Argentina, $0.7057\text{--}0.7115$ and $8.3\text{--}10.7\%$)²⁻⁶. These isotopic ratios, and the overall positive correlation of $^{87}\text{Sr}/^{86}\text{Sr}$ with $^{18}\text{O}/^{16}\text{O}$ ratios observed for the lavas of the CVZ, form the basis of various crustal contamination and assimilation models²⁻⁸, in which the contrasted thicknesses of continental crust beneath the NVZ ($\sim 45 \text{ km}$) and the CVZ ($\sim 70 \text{ km}$) have an essential role^{9,10}.

Sr- and O-isotope data for eight active volcanic centres in the SVZ of Chile are presented in Table 1, which shows the following features: (1) $^{87}\text{Sr}/^{86}\text{Sr}$ ratios range from 0.7036 to 0.7043 and show a small, but significant, increase southwards (from $36^\circ 50'\text{S}$ to $41^\circ 20'\text{S}$) and eastwards (from $71^\circ 25'\text{W}$ to $72^\circ 37'\text{W}$); (2) $\delta^{18}\text{O}$ values range from 5.2 to 6.8%, although it is not possible to demonstrate a significant geographical variation; (3) samples of different chemical compositions from individual centres exhibit little or no significant differences in either $^{87}\text{Sr}/^{86}\text{Sr}$ or $^{18}\text{O}/^{16}\text{O}$ isotopic ratios. This is most clearly demonstrated by Villarica volcanic centre, where basalt, andesite, and rhyolite ($69.2\% \text{ SiO}_2$) respectively yield $^{87}\text{Sr}/^{86}\text{Sr}$ values of 0.70397 ± 4 (2σ error), 0.70397 ± 4 , and 0.70389 ± 6 , together with $\delta^{18}\text{O}$ values of 5.8, 5.8, and 6.3%.

The $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{18}\text{O}/^{16}\text{O}$ ratios for the volcanic centres of the SVZ are the most primitive yet measured for the young rocks of the Andean Cordillera, and fall within the overall range of $^{87}\text{Sr}/^{86}\text{Sr} = 0.7040 \pm 0.0005$ and $\delta^{18}\text{O} = 6.0 \pm 0.5\%$, considered representative of 'normal' mantle compositions. Similar $^{87}\text{Sr}/^{86}\text{Sr}$ ratios within the narrow range 0.7041 ± 0.0002 have been reported^{11,12} for basalts and dacites from Recent volcanoes between 36°S and 46°S , and values of $0.7038\text{--}0.7039$ for basalts from Antuco¹². Whether the small variations observed in the isotopic ratios in the present work are due to sub-crustal source heterogeneity or to a very small degree of localized and variable interaction of these mantle-derived mag-

Table 1 Variation of $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}\%$ with latitude and longitude in eight volcanic centres in the southern Andean volcanic zone of central Chile

Lat. S	Long. W		2σ error	2σ error
36 50'	71 25'	N. de Chillan	□	□
37 25'	71 17'	Antuco	○▲	▲○
38 24'	71 34'	Lonquimay	○△	△○
38 43'	71 44'	Llaima	△△○	△
39 25'	71 57'	Villarica	★▲	▲★
39 56'	72 03'	Choshuenco	○△	○△
41 07'	72 30'	Osorno	▲▲	▲
41 20'	72 37'	Calbuco	▲○	○▲

$^{87}\text{Sr}/^{86}\text{Sr}$ $\delta^{18}\text{O}(\text{‰SMOW})$

▲, Basalt; △, basaltic andesite; ○, andesite; □, dacite; ★, rhyolite.

Table 2 Variation of $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}\%$ with latitude and longitude for five volcanic centres in the central Andean volcanic zone of northern Chile, northwestern Argentina and southwestern Bolivia

Lat. S	Long. W		2σ error	2σ error
24 14'	66 33'	San Jeronimo (NW Argentina)	△○	△○
24 05' -24 30'	68 10' -68 37'	El Negrillar, (N Chile)	△○	△○
23 22'	67 44'	Lascar, (N Chile)	○□	○□
22 26'	67 55'	Tocorpuri, (N Chile)	★□	★□
21 55'	66 55'	Sierra de Lipez, (SW Bolivia)	□★	★□

$^{87}\text{Sr}/^{86}\text{Sr}$ $\delta^{18}\text{O}(\text{‰SMOW})$

△, Basaltic andesite; ○, andesite; □, dacite; ★, rhyolite; I, ignimbrite.

mas with lower crust, possibly as a result of increasing crustal thickness away from the subduction zone, cannot be ascertained on the basis of only the Sr- and O-isotope evidence. That the small $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}$ variations observed do not change with differentiation (as reflected by SiO_2 content), clearly indicates that the magmas of the SVZ have evolved free of interaction with crustal rocks that resided at the Earth's surface. These data closely mirror the Sr- and O-isotope relationships reported recently in a plutonic rock sequence which resulted solely from crystal fractionation of a mantle-derived magma in a closed system, forming part of an oceanic-arc complex on Guadalcanal in the Solomon Islands¹³. Moreover, there is no indication whatever that SVZ magmas were affected by a process of combined fractional crystallization-assimilation, such as that modelled by Taylor¹⁴ and invoked by James⁸ for the Barroso and Arequipa volcanics in southern Peru, during transit and subsequent high-level differentiation before eruption.

Sr- and O-isotope data for five volcanic centres in the CVZ are presented in Table 2. The following points may be noted: (1) $^{87}\text{Sr}/^{86}\text{Sr}$ ratios for all centres except Sierra de Lipez fall in the range 0.7056–0.7096, with small, but significant, within-group variations. (2) $\delta^{18}\text{O}$ values from the same four centres range from 7.3 to 9.6‰, with individual centres also showing significant within-group variations, which are roughly correlated with $^{87}\text{Sr}/^{86}\text{Sr}$ variations. (3) Sierra de Lipez has wider ranges in $^{87}\text{Sr}/^{86}\text{Sr}$ from 0.7090 to 0.7149, and in $\delta^{18}\text{O}$ from 8.0 to 10.8‰, which do not correlate directly with bulk composition. (4) The Lascar Centre is the only volcanic centre studied for which there is any significant correlation between Sr- and O-isotopic composition and SiO_2 content (that is bulk composition).

In addition, detailed consideration of chemical and isotopic data (in preparation) shows a positive correlation between $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}$ with Rb p.p.m., as well as negative

hyperbolic correlation between $^{87}\text{Sr}/^{86}\text{Sr}$ p.p.m. for Lascar and, particularly, for the Sierra de Lipez centre. The latter phenomenon results from contamination with crustal Sr during high-level fractional crystallization of a range of subsequently erupted magmas, as described in the model of Francis *et al.*³ for Cerro Galan in north-west Argentina. There is also a pronounced positive correlation between initial $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{87}\text{Rb}/^{86}\text{Sr}$ for Sierra de Lipez, the resulting 'pseudo-isochron' being interpreted as a mixing line with no age significance^{3,7}.

In agreement with previous measurements in the CVZ^{2-5,8,15}, none of the Sr- and O-isotope data shown in Table 2 fall within the range of 'normal' mantle compositions. All five CVZ volcanic centres investigated here are enriched in both ^{87}Sr and ^{18}O relative to 'normal' mantle, as well as to the volcanic centres of the NVZ in Ecuador and the SVZ in southern Chile.

Numerous petrogenetic models have been discussed to account for isotopic trends of the CVZ, as well as other isotopic data based on Nd (ref. 16) and Pb (ref. 17). The models which we consider least satisfactory for explaining the major and systematic isotopic and geochemical differences between the CVZ on the one hand, and the NVZ and SVZ on the other, as well as the intra-CVZ variation, are: (1) bulk melting of subducted sediments^{4,15}; (2) major, localized sub-CVZ mantle heterogeneity, enhanced by input of marine Sr ($^{87}\text{Sr}/^{86}\text{Sr} \sim 0.709$) from the subducted lithosphere¹⁶; (3) large-scale crustal anatexis of deep CVZ crust. Substantial arguments against models of this type have been presented elsewhere (see, for example, refs 2, 3, 5-10, 19). In addition, there are severe experimentally-derived constraints on large-scale bulk melting of continental crust to produce massive volumes of calc-alkaline magmas of intermediate composition (see, for example, refs 20, 21).

More acceptable models have centred around bulk or selective crustal contamination of mantle and/or subduction zone derived magmas, which may have started with characteristic mantle-type $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}$ values. Contamination processes within the deep continental crust by the interaction of a mantle-derived melt with Mesozoic or older igneous rocks in the lower crust, which has been thickened by plutonic activity, can produce the enhancement of $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{18}\text{O}/^{16}\text{O}$ ratios into the ~ 0.705 - 0.708 and $\delta^{18}\text{O} \sim 7$ - 9% ranges commonly observed in the CVZ. Subsequent high-level plagioclase fractionation in some volcanic centres accompanied by further crustal contamination may then lead to the isotopic patterns observed at Cerro Galan³, Lascar and Sierra de Lipez (Table 2, and in preparation). It is possible that deep crustal contamination of mantle-derived magmas is dominated by bulk assimilation of basement rocks, or, more likely, in view of the absence of crustal xenoliths, of partial melts thereof, whilst high-level crustal contamination of a plagioclase-fractionating magma may be dominated by selective contamination processes (see refs 3, 6, 17), and/or by combined assimilation and fractional crystallization^{7,14}. We find arguments (see refs 8, 16, 22) for the existence of a sub-CVZ mantle-type source region with $^{87}\text{Sr}/^{86}\text{Sr} \sim 0.705$ - 0.707 , and corresponding arguments against crustal contamination^{16,22}, unpersuasive. In this regard, a single sample of undersaturated alkali-basalt (collected by R. S. Thorpe and P. W. Francis) from a recent Tuff-ring at Salar de Uyuni in the CVZ of western Bolivia, close to the crustal keel of the Andes some 300 km from the Chile trench, has a $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.70400 ± 7 , within the typical range of all analysed NVZ and SVZ lavas. This may be representative of sub-CVZ and sub-Andean mantle^{3,9}. Andesite lavas from the same area have ratios in the range 0.7053 - 0.7063 , suggestive of crustal contamination.

The comparative data for the SVZ and CVZ of the Andes, reported here, lead us to agree with Harmon *et al.*⁵ that magmatism throughout the length of the Andean Cordillera is mantle-derived, and that it is the slow rise of mantle-derived basic and intermediate magmas through the exceptionally thick (~ 60 - 70 km) crustal keel of the Central Andes that provides an extended opportunity for crustal interaction before eruption.

In the SVZ and NVZ crustal interaction is minimal, allowing mantle-derived magmas to rise largely unmodified into the crust, where they differentiate before eruption without modifying their isotopic composition. In contrast, mantle-derived magmas in the CVZ undergo early interaction with plutonic rocks in the lower crust involving anatexis or bulk partial melting and homogenization before transit through the upper crust, where they may undergo additional contamination due to high-level fractional crystallization-assimilation processes. We do not mean to understate the importance of lower crustal involvement such as that recently advocated by Hawksworth *et al.*¹⁸ for the Purico volcanic centre in the CVZ of northern Chile, but rather wish to shift the emphasis to stress the predominant mantle role in Andean magmatism.

Isotopic work was carried out at Oxford University and at the Scottish Universities Research and Reactor Centre, East Kilbride, supported by grants from the Natural Environment Research Council. We thank R. Goodwin, M. Humm and J. Borthwick for technical assistance, and R. S. Thorpe for critical comments.

Received 9 August 1982; accepted 25 February 1983.

- Cummings, D. & Schiller, G. I. *Earth-Sci. Rev.* **7**, 97-125 (1971).
- Francis, P. W., Moorbath, S. & Thorpe, R. S. *Earth planet. Sci. Lett.* **37**, 197-202 (1977).
- Francis, P. W., Thorpe, R. S., Moorbath, S., Kretzschmar, G. A. & Hammill, M. *Earth planet. Sci. Lett.* **48**, 257-267 (1980).
- Magaritz, M., Whitford, D. J. & James, D. E. *Earth planet. Sci. Lett.* **40**, 220-230 (1978).
- Harmon, R. S., Thorpe, R. S. & Francis, P. W. *Nature* **290**, 396-399 (1981).
- Briqueu, L. & Lancelot, J. R. *Earth planet. Sci. Lett.* **43**, 385-396 (1979).
- De Paolo, D. J. *Earth planet. Sci. Lett.* **53**, 189-202 (1981).
- James, D. E. *Earth planet. Sci. Lett.* **57**, 47-62 (1982).
- Thorpe, R. S. & Francis, P. W. *Tectonophysics* **57**, 53-70 (1979).
- Coulon, C. & Thorpe, R. S. *Tectonophysics* **77**, 79-93 (1981).
- Stern, C. R. & Futa, K. *EOS* **63**, 456 (1982).
- Hickey, R. L., Frey, F. A., Lopez-Escobar, L. & Munizaga, F. *Geol. Soc. Am. Abstr.* (1982).
- Chivas, A. R., Andrews, A. S., Sinha, A. K. & O'Neil, J. R. *Nature* **300**, 139-143 (1982).
- Taylor, H. P. *Earth planet. Sci. Lett.* **47**, 243-254 (1980).
- Klerkx, J., Deutsch, S., Pichler, H. & Zeil, W. *J. Volcan. geotherm. Res.* **2**, 49-71 (1977).
- Hawksworth, C. J. *et al. Earth planet. Sci. Lett.* **42**, 45-57 (1979).
- Tilton, G. R. & Barreiro, B. A. *Science* **210**, 1245-1247 (1980).
- Hawksworth, C. J., Hammill, M., Gledhill, A. R., Calst ren, P. & Rogers, G. *Earth planet. Sci. Lett.* **58**, 240-254 (1982).
- Thorpe, R. S. & Francis, P. S. *Origin of Granite Batholiths—Geochemical Evidence* (eds Atherton, M. P. & Tarney, J.) 65-75 (Shiva, Orpington, 1979).
- Wyllie, P. J. *Tectonophysics* **43**, 41-71 (1977).
- Wyllie, P. J. *Geol. Rdsch.* **70**, 128-153 (1981).
- Hawksworth, C. J. *Origin of Granite Batholiths—Geochemical Evidence* (eds Atherton, M. P. & Tarney, J.) 76-89 (Shiva, Orpington, 1979).

Cation-ratio dating of petroglyphs from the Western Great Basin, North America

Ronald I. Dorn* & David S. Whitley†

* Department of Geography and † Institute of Archaeology, University of California, Los Angeles, California 90024, USA

Using a new age-determination technique for rock varnish, cation-ratio dating¹, we have obtained absolute dates for varnishes within the pecked surfaces of five petroglyphs from the Coso Range, Inyo County, California. Our age determinations, based on analytical data gathered directly from petroglyphs, revise previously proposed chronologies and suggest a greater antiquity for Great Basin rock art in the western United States.

Petroglyphs in arid and semi-arid lands are frequently etched into a dark coating called rock varnish, composed of oxides and hydroxides of manganese and iron, clay minerals, and minor and trace elements that accrete on rock surfaces^{2,3}. After pecking the petroglyph element through the varnish, the newly exposed rock is subject to the subaerial accretion of varnish constituents. Simultaneous with the onset of revarnishing, but at a slower rate, cation-exchange complexes in the illite and montmorillonite clays⁴ and oxides⁵ that are dominant in varnish are subject to leaching.

Cation-ratio dating is based on the concept that certain cations (such as Na, K, Ca, P and Mg) are more mobile than Ti (refs 6-8), and a ratio of one or more of these mobile cations to Ti decreases with time. Any base can be used in a ratio with

Ti. However, K and Ca are used in this study, because the analytical method used (particle induced X-ray emission or PIXE)⁹ gives the most reliable ratio for K+Ca:Ti. Ti is used, rather than Al or Fe, because it does not appear to be influenced by manganese-concentrating microorganisms found on varnish³, it is present in readily measurable quantities, and it is extremely immobile in semi-arid and arid environments⁶⁻⁸.

Several assumptions are involved in cation-ratio dating. (1) The initial ambient K+Ca:Ti ratio has remained relatively constant over the period of varnish accretion. (2) Varnish serves as a cation-exchange complex, and the exchange capacity is similar for compared samples. (3) Bases such as K and Ca are more mobile than Ti in varnish. (4) The rates of reduction of the cation ratio (in this case K+Ca:Ti) are similar for compared samples. These assumptions, which have been tentatively verified for the Coso region, and other potential limitations of cation-ratio dating are discussed in ref. 1.

The age-determination of varnish cation ratios is based on their calibration to independently dated surfaces in the Coso volcanic field and vicinity, southern California. Rock varnish on 12 Coso volcanics, dated from 39,000 to more than 3 Myr by the K-Ar method¹⁰, was sampled in the field. Varnish was removed from the underlying rock by careful scraping with a tungsten carbide needle under $\times 10$ –45 stereomagnification and then analysed by the PIXE technique⁹. Each cation ratio of varnish on the Coso volcanics, plotted in Fig. 1, represents an average of five PIXE analyses.

The $\sim 10,500$ -yr-old high lake level of nearby Searles Lake^{11,12} was used to extend the calibration of varnish cation-ratios. The clay-sized fraction of soils from 16 sites in the Coso Range was selected to represent the modern K+Ca:Ti ratio, because the detritus incorporated into varnish is probably derived from soil deflation^{3,13}, and the particle size incorporated into varnish is usually $< 2 \mu\text{m}$ (refs 1, 3).

Figure 1 presents the empirical cation-leaching curve for the Coso-Searles Lake region. Further details on the construction of this preliminary estimate of varnish-leaching rates in the Coso region are presented in ref. 1. The lines in Fig. 1 can be described by the following equation:

$$y = \begin{cases} 7.468 - 0.1701 \log x & \text{for } 10^2 < x < 1.05 \times 10^4 \\ 8.211 - 2.433 \log x & \text{for } 1.05 \times 10^4 < x < 1.50 \times 10^5 \\ 5.724 - 1.296 \log x & \text{for } 1.50 \times 10^5 < x < 10^7 \end{cases}$$

where x is in yr BP and y is the K+Ca:Ti varnish cation ratio.

The Environmental Branch, Naval Weapons Center China Lake, granted us permission to collect five petroglyph varnish samples from the Coso Range. These petroglyphs are located at the Wheeler Prospect and Birchim Springs sites (Fig. 2), which are representative of the Coso Range Petroglyph assemblage¹⁴. The specific petroglyphs (Fig. 3) were selected on the basis of maximum observed varnishing for an individual style and similar microenvironmental conditions.

Scraping varnish from within the petroglyphs was done in the field. In the laboratory fragments of the underlying rock that were incorporated into the sample were extracted under $\times 45$ stereomagnification. The remainder, petroglyph varnish, was processed like natural varnish samples for PIXE analysis¹. The average and standard deviation cation ratios of four of the petroglyphs were determined by PIXE analyses of three separate samples from each petroglyph. There was only enough material in petroglyph Fig. 3A for one PIXE analysis. Hence, the uncertainty reported here for Fig. 3A is based on the experimental error of the single PIXE run. The accuracy of the date for Fig. 3A is therefore suspect.

The average cation ratio of each petroglyph was compared with the cation-leaching curve in Fig. 1, and an absolute age determination was obtained on the X-axis. The upper and lower standard deviations of the PIXE analyses (or experimental uncertainty in the case of Fig. 3A) provide an error margin (EM) for the age estimates. For example, Fig. 3D has an average varnish K+Ca:Ti ratio of 6.10. This corresponds to an age of 6,400 yr BP on Fig. 1. The upper standard deviation

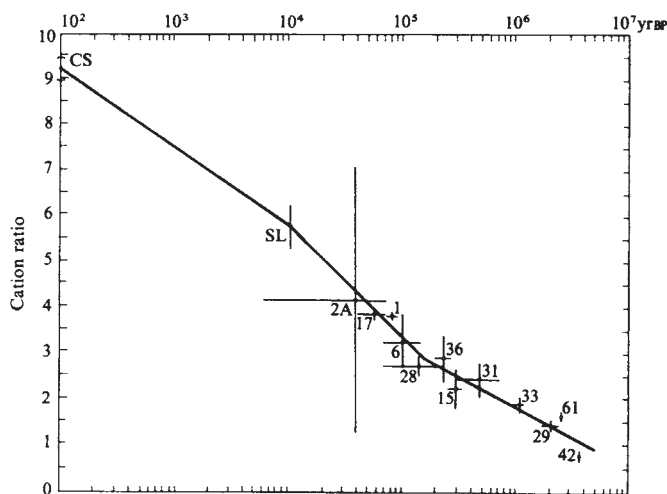


Fig. 1 Varnish cation-leaching curve for the Coso region. The cation ratio used is K+Ca:Ti. The numbers represent K-Ar dated volcanic surfaces in the Coso Range, identified in ref. 19. The horizontal and vertical bars reflect the K-Ar age uncertainties and cation ratio standard deviations, respectively. SL, Varnish ratio for the 'Searles Lake' late Pleistocene high stand. CS, clay-sized cation ratio of 16 Coso soils, collected near the analysed volcanics; the brackets on the left-hand margin of the graph represent the standard deviation of the soil cation ratio. The lines from SL to 36 and from 36 to 42 are least-squares regression lines with correlation coefficients of -0.99 and -0.98 , respectively. The line from SL to CS is drawn between end points.

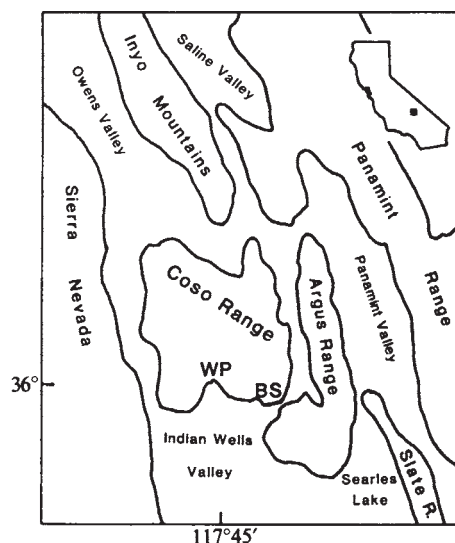


Fig. 2 Map showing the physiography of the Coso Range and vicinity and the location of the Wheeler Prospect (WP) and Birchim Springs (BS) petroglyph sampling sites, Inyo County, California.

($+0.21$ or a ratio of 6.31) has an equivalent age of 5,600 yr BP, and the lower standard deviation (5.89) is assigned 8,600 yr BP. Therefore, the estimated age of Fig. 3D is 6,400 yr BP, with a substantial error margin of 5600–8600 yr BP. Figure 3A is estimated to be 580 yr old (EM 120–2,900 yr BP). Figure 3B corresponds to 4,300 yr BP (EM 2,200–8,600 yr BP). Figure 3C dates to 6,000 yr BP (EM 2,250–13,500 yr BP), and Fig. 3E is equivalent to 1,700 yr BP (EM 1,350–2,050 yr BP). Note that while the error margins are quite large, these dates are the only absolute age determinations available for North American petroglyphs that are based on an objective dating method.

Previous rock art research in the Great Basin established relative and absolute chronologies. The relative ordering, based on stylistic and superimpositional analyses, presents Great Basin curvilinear abstract as oldest, followed successively by

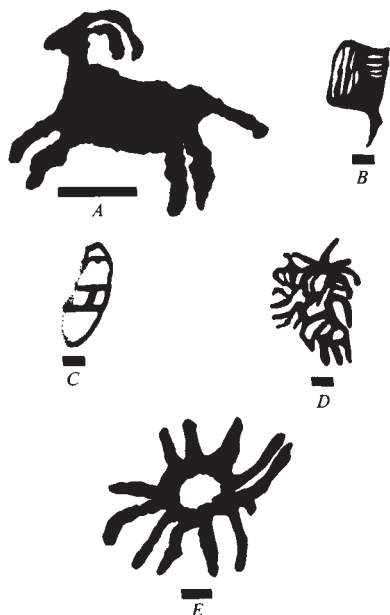


Fig. 3 Petroglyph motifs in the Coso Range dated by rock varnish cation ratios. Scale, 5 cm. *A*, Great Basin representational bighorn side-facing sheep, Wheeler Prospect site, petroglyph CS 2. *B*, Great Basin rectilinear abstract, Wheeler Prospect site, petroglyph CS 3. *C*, Great Basin curvilinear abstract, Wheeler Prospect site, petroglyph CS 4. *D*, Great Basin curvilinear abstract, Birchim Springs site, petroglyph BSS2. *E*, Great Basin curvilinear abstract, Wheeler Prospect site, petroglyph Cs 1.

Great Basin rectilinear abstract and Great Basin representational styles¹⁵⁻¹⁷. Heizer and Baumhoff's (ref. 16, p. 284) absolute chronology suggested the following dates for the above respective styles: ~1000 BC to AD 1500, AD 1 to AD 1500, and AD 1 to AD 1500. Within the Coso Range, Grant (ref. 14, p. 58) argued that the Great Basin representational style had an extended span of ~1000 BC to AD 1000, and that the production of all styles began within the past 3,000–4,000 yr.

However, Heizer and Baumhoff's (ref. 16, p. 284) absolute chronology is based on two assumptions: that the youngest petroglyphs in the Great Basin are 150–200 yr old and that petroglyphs with the best developed varnish are 15–20 times older than the youngest motifs. No data were presented to support these assertions. Similarly, Grant's view¹⁴ that petroglyph production began 3,000–4,000 yr ago in the Coso Range is based on the belief that varnish erodes in periods of accelerated rainfall, the last of which Grant felt occurred about 3,000–4,000 yr BP. Grant's view¹⁴ of varnish erosion in pluvial periods has been criticized recently³ and we feel that the assumptions of both proposed chronologies are unsubstantiated, serving to underline the need for a chronology based on more objective measures.

While any interpretation of our petroglyph dates is provisional, given the small sample of dates and the wide error margins, some preliminary observations are in order. First, the cation-ratio dates suggest that rock art production within the Great Basin is almost twice as old as previously hypothesized. Using only the conservative margin of 5,600 yr BP for Fig. 3D (average age of 6,400 yr BP; EM 5,600–8,600 yr BP), the curvilinear abstract motif is at least 2,600 yr older than the maximum age of 3,000 yr suggested by Heizer and Baumhoff¹⁶, and 2,600–1,600 yr older than the 3,000–4,000 yr BP date suggested by Grant¹⁴.

Second, using the conservative younger error margin of Fig. 3D (5,600 yr BP) and the older error margin of Fig. 3E (2,050 yr BP), the Great Basin curvilinear abstract style elements have a temporal range of ~3,600 yr. However, the temporal range of the curvilinear abstract style is probably greater than this estimate.

Third, our dates provide weak support for the previously hypothesized relative stylistic chronology (see ref. 16). The

average age estimates of the curvilinear style elements (Fig. 3C, D, E) and the rectilinear abstract style element (Fig. 3B) overlap. This may indicate that production of these two styles was in part coterminous, as originally hypothesized by Heizer and Baumhoff¹⁶. Tentatively, the average cation-ratio dates of Fig. 3D (6,400 yr BP) and Fig. 3C (6,000 yr BP) suggest a greater antiquity for the initiation of the curvilinear style than the average date of the rectilinear motif (Fig. 3B; 4,300 yr BP). Verification of this hypothesis will require further dating. Similarly, the representational motif (Fig. 3A) appears to be the youngest petroglyph sampled. However, the older error margin of Fig. 3A overlaps with the youngest error margin of Fig. 3B. Our reinterpretation of Bard's¹⁸ neutron activation data tables of petroglyphs from Grimes Point, Fallon Nevada, using the Na+Ca:Ti cation ratio, indicates that the pit-and-groove style (see ref. 16) is older (more leached) than Great Basin curvilinear style motifs sampled at the Grimes site.

Cation-ratio dating is not limited to dating petroglyphs in the Coso Range. It may be used to verify or correct a relative chronologic sequence at any petroglyph site. With surfaces of known age nearby, it may be possible to develop a separate cation-leaching curve to make absolute age-determinations. Cation-ratio dating may also be used in the relative and absolute dating of surface artefacts, in arid geomorphology, and palaeoenvironmental research in arid and semi-arid regions.

NSF grant SES80-24555 to Dr T. M. Oberlander supported the development of the Coso cation-leaching curve. The petroglyph dating was funded in part by grants from the University of California. We thank Tom Dodson, Denise Dorn, Glen Roqmore, Ted Oberlander, Carolyn Shepherd and Theresa Whitley for field assistance.

Received 19 October 1982; accepted 21 February 1983.

1. Dorn, R. I. *Quat. Res.* (in the press).
2. Dorn, R. I. & Oberlander, T. M. *Z. Geomorphol. N. F.* **25**, 420–436 (1981).
3. Dorn, R. I. & Oberlander, T. M. *Prog. Phys. Geog.* **6**, 317–367 (1982).
4. Potter, R. & Rossman, G. R. *Science* **196**, 1446–1448 (1977).
5. Jenne, E. A. in *Trace Inorganics in Water* (ed. Gould, R. F.) 337–387 (American Chemical Society, Washington DC, 1968).
6. Marchand, D. E. *U.S. Geol. Surv. Prof. Pap.* **352-J**, 379–424 (1974).
7. Gardner, L. R. *Chem. Geol.* **30**, 151–165 (1980).
8. Colman, S. M. *U.S. Geol. Surv. Prof. Pap.* No. 1246 (1982).
9. Cahill, T. A. in *Short-lived Radionuclides in Chemistry and Biology* (eds Root, J. W. & Krohn, K. A.) 511–522 (American Chemical Society, Washington DC, 1981).
10. Duffield, W. A., Bacon, C. R. & Dalrymple, C. B. *J. geophys. Res.* **85**, 2381–2404 (1980).
11. Smith, G. I. in *Paleolimnology of Lake Biwa and the Japanese Pleistocene* Vol. 4 (ed. Horie S.) 577–604 (Kyoto University Press, 1976).
12. Smith, G. I. *U.S. Geol. Surv. Prof. Pap.* No. 1043 (1979).
13. Péwé, T. L., Péwé, E. A., Péwé, R. H., Journauz, A. & Slatt, R. M. *Geol. Soc. Am. Spec. Pap.* **186**, 169–198 (1981).
14. Grant, C. *Rock Drawings of the Coso Range, Inyo County, California* (Maturango Museum, China Lake, 1968).
15. Steward, J. H. *Univ. Calif. Publ. Am. Archaeol. Ethnol.* **24**, 47–238 (1929).
16. Heizer, R. F. & Baumhoff, M. A. *Prehistoric Rock Art of Nevada and Eastern California* (University of California Press, Berkeley, 1962).
17. von Werthof, J. C. *Univ. Calif. Archaeol. Surv. Rep.* 65 (1965).
18. Bard, J. C. thesis, Univ. California, Berkeley (1979).
19. Duffield, W. A. & Bacon, C. R. *U.S. Geol. Surv. Miscellaneous Investigation Map I-1200* (1981).

Evolutionary links between apposition and superposition optics in crustacean eyes

D.-E. Nilsson

Department of Zoology, University of Lund,
S-223 62 Lund, Sweden

The two fundamental principles of image formation in compound eyes, apposition and superposition, have long been known to exist in both insects and crustaceans¹. The scattered distribution of the two optical mechanisms²⁻⁴ indicates that one type can evolve into the other, but as they are so fundamentally different, it has been an enigma how such evolution could occur. In crustaceans, superposition eyes are present among the orders Mysidacea, Euphausiacea and Decapoda⁵. I now report that

the planktonic larvae of most euphausiids and decapods possess a transparent type of apposition eye, designed for planktonic life, that is pre-adapted for superposition optics. This discovery suggests that the evolution of superposition eyes in crustaceans has its origin in the special apposition optics found in the larval eyes, thus providing the first evidence of an evolutionary link between apposition and superposition optics.

The two basic types of arthropod compound eye, apposition and superposition, use widely different optical principles^{1,5}. The apposition eye consists of optically isolated units (ommatidia), each having a lens forming an inverted image (Fig. 1a). In superposition eyes, the lens systems of many ommatidia cooperate to form a superimposed erect image. The crucial feature of a superposition eye is for the lens system in each ommatidium to direct light towards the retina (rhabdom layer), at the same side of the ommatidial axis as it enters the eye (see Fig. 1). This can be achieved either by refraction in a lens-cylinder gradient^{1,6} (Fig. 1b) or by reflections in an orthogonal system of mirrors⁷⁻⁹ (Fig. 1c), resulting in the distinction between refracting and reflecting superposition eyes.

Both apposition and superposition eyes are present in crustaceans as well as in insects²⁻⁴ but the reflecting superposition type is reported for crustaceans alone^{5,9}. In crustaceans, superposition eyes exist in mysids, euphausiids and decapods but apposition eyes also occur within these groups^{4,5} (Table 1).

The apposition principle is a simple and straightforward one that can be traced back, in its most primitive form, to annelid worms³. The superposition principle is more complex, and superior to apposition due to the improved light-gathering capacity³. Based on this argument and on the scattered occurrence of superposition eyes, it seems reasonable to believe that superposition eyes have been derived from apposition eyes. However, the problem has been to understand the course of such an evolution. No hypotheses have been presented in this respect, because no types intermediate between the two discrete principles have been found. Also, due to the conflict between forming erect and inverted images, it was hard to imagine the design of functional intermediate types. The investigation, presented here, of the larval apposition eyes of euphausiids and decapods partly solves the problem by revealing functional links between apposition and superposition.

For an optical investigation, the furcilia III larva of the euphausiid *Thysanoessa raschii* (M. Sars) and the eleventh-stage larva (megalopa) of the decapod prawn *Macrobrachium rosenbergii* (de Man) were chosen. These species have, as adults,

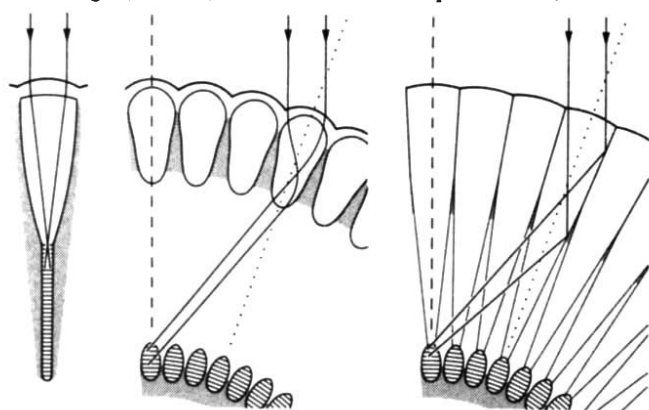


Fig. 1 Schematic drawings of the optical systems in compound eyes. In the ommatidium of the apposition eye (a), the crystalline cone is contiguous with the rhabdom (striped area) and parallel light is focused on the rhabdom tip. Due to the pigment screen (shaded area), the ommatidia are optically isolated from their neighbours. In the superposition eyes (b, c), the clear zone distal to the rhabdoms allows light to pass across the ommatidia to form superposition images, either by refraction (b) or reflection (c). Parallel rays incident obliquely on ommatidia in b and c are projected on the rhabdom of the ommatidium which is aligned with the light source (broken line).

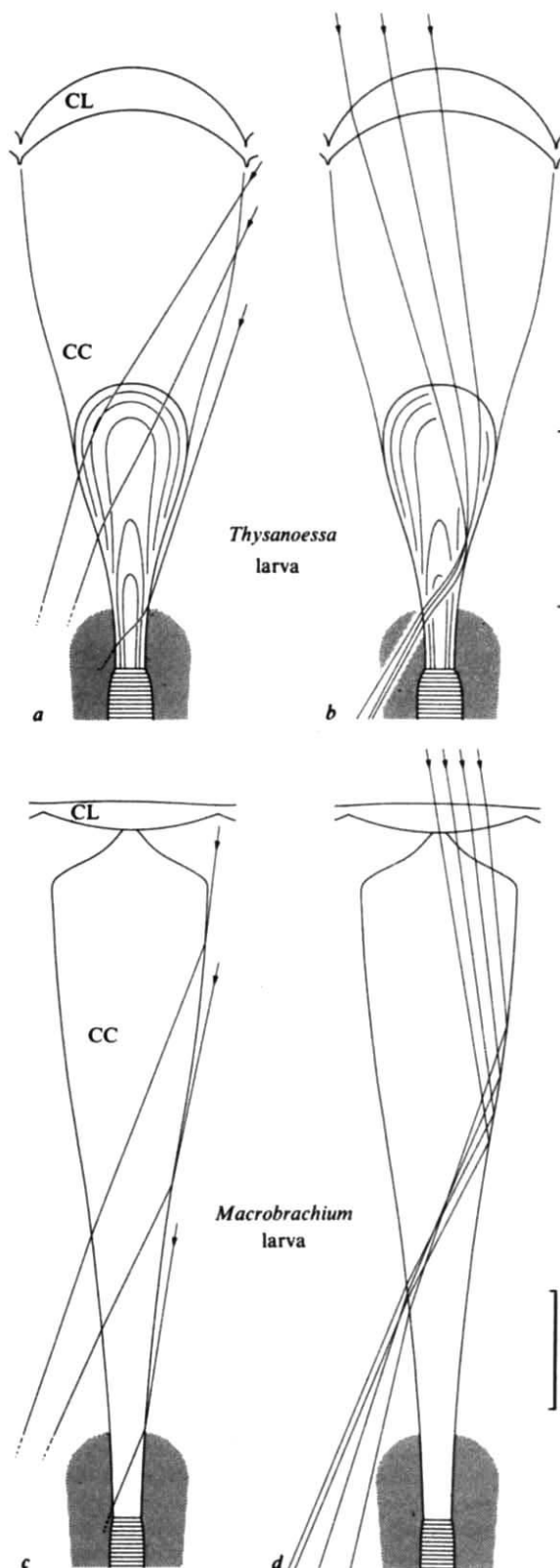


Fig. 2 Ray-tracing in ommatidial models of *Thysanoessa* and *Macrobrachium* larvae. Optical isolation (a, c) is demonstrated by side-entering rays. The formation of superposition rays, incident at 9° to the axes, is shown (b, d) (absorption in screening pigment disregarded). The proximal gradient in the crystalline cone of *Thysanoessa* (a, b) is represented by the following isoindex lines: 1.37, 1.41, 1.45, 1.49, 1.53 and 1.57, from the distal periphery to the proximal centre. The crystalline cone of *Macrobrachium* (c, d) is almost homogeneous ($n = 1.396$, with a slight increase to 1.405 in the proximal tip) and surrounded by extracellular fluid ($n = 1.339$)¹⁵. CL, Corneal lens; CC, crystalline cone; shaded area, screening pigment; striped area, rhabdom tips. Scale bars, 20 μm .

Table 1 Distribution of eye types in mysids, euphausiids and decapods

Group	Eye type
Mysidacea	Refracting superposition
Euphausiacea adults	Refracting superposition
Euphausiacea larvae*	Apposition
Decapoda adults:	
Natantia	Reflecting superposition
Macrura	Reflecting superposition
Anomura (some)	Reflecting superposition
Anomura (other)	Apposition
Brachyura	Apposition
Decapoda larvae	Apposition

* Late instars have refracting superposition eyes.

refracting and reflecting superposition eyes respectively, but the planktonic larvae have 'transparent' apposition eyes without any pigment screen around the crystalline cones. For refractive index measurements, fresh crystalline cones were isolated in a salt solution of 420 mM NaCl, 25 mM MgCl₂, 12 mM CaCl₂, 10 mM KCl and 10 mM HEPES buffer. Microinterferometry was applied and the interference patterns of intact cones were analysed by the method of Nilsson *et al.*¹⁰ to determine the refractive index distribution. The focal length of the corneal lenses was measured on clean pieces of cornea, according to the method of Bryceson¹¹. Based on these data, ommatidial models were constructed and analysed by ray-tracing¹² (Fig. 2).

The euphausiid larva displayed a refractive index gradient in the proximal part of the crystalline cone. An image is focused within this gradient, about half-way along its length, and the proximal part of it acts as a short graded-index light-guide, propagating light from its own corneal lens to the rhabdom. Rays from adjacent corneal lenses are rejected by the proximal gradient (Fig. 2a), thus compensating for the lack of screening pigment along the crystalline cones. This optical system is almost identical with that found in some pelagic amphipods where the transparency achieved subserves the camouflage¹³. Although it is an apposition eye (having optically isolated ommatidia and lacking the clear zone), superposition rays will be formed as an unused secondary effect of the proximal gradient (Fig. 2b). The resulting afocal ray-path is in principle that of a refracting superposition eye (Fig. 1), which is realized in the adult.

In the larval eye of the decapod prawn, a distant source is focused on the rhabdom tip (as in a 'conventional' apposition eye), but also in this case, optical isolation is maintained without screening pigment between the cones. This is achieved simply by a high refractive index in the crystalline cones which deflects light, entering from adjacent facets, away from the rhabdom (Fig. 2c). The same mechanism was recently reported to result in less discernible eyes in the isopod *Astacilla*¹⁴. As a result of the high refractive index, total internal reflections will occur (Fig. 2d) which provide the basis for reflecting superposition optics (Fig. 1) (which is known to exist in the adult¹⁵), but as the ommatidia are optically isolated (apposition type) in the larva, superposition rays cannot be used for vision.

Histological sections from several other euphausiid and decapod larvae were examined by light microscopy. Due to the absence of screening pigment between the crystalline cones, all eyes were superficially similar. However, two groups could be distinguished, closely resembling either the *Thysanoessa* type with a dense rounded structure in the proximal part of the cone (representing the gradient in Fig. 2a, b) or the *Macrobrachium* type lacking this structure. Table 2 suggests that the optical findings in the *Thysanoessa* and *Macrobrachium* larvae are general for euphausiid and decapod larvae, respectively. However, among euphausiid larvae there is one known exception—*Thysanopoda tricuspidata*—which has superposition eyes from the first furcilia instar¹⁶.

Planktonic organisms are usually unpigmented because this is the only means of camouflage in their habitat. Therefore, the most probable reason for the special optical design of the

Table 2 Morphological classification of larval eyes in euphausiids and decapods

Species	Eye type
Euphausiacea	
<i>Meganyctiphanes norvegica</i>	T
<i>Stylocheiron</i> sp.	T
<i>Euphausia</i> sp.	T
Decapoda	
<i>Sergestes</i> sp.	M
<i>Homarus vulgaris</i>	M
<i>Galathea strigosa</i>	M
<i>Pagurus</i> sp.*	M
<i>Carcinus maenas</i> *	M

All observations are from histological sections. T, *Thysanoessa* type; M, *Macrobrachium* type. The euphausiids are early furcilia instars. The five decapod species correspond to the five decapod groups in Table 1. Mysids are not included due to the lack of free-living larval stages.

* Species with apposition eyes also in the adult stage.

apposition eyes of euphausiid and decapod larvae is to remove the need for a pigment screen, and hence to make the eyes maximally transparent. The possibility that the larval eye types result from the presence of superposition eyes in the adults is ruled out by the fact that crabs, which have apposition eyes as adults, have the same type of larval eye as *Macrobrachium* and that the transparent eye type has been found in various other crustaceans^{13,14,17,18}. To achieve transparent eyes, euphausiid and decapod larvae use optical mechanisms which provide the functional bases for the superposition principles used by the adults. The most important changes required for the larval eyes to convert to superposition type are the creation of a clear zone^{2,5} distal to the rhabdom layer and for the reflecting type, a transformation from hexagonal to square facets⁹.

The discovery of apposition eyes being pre-adapted for superposition optics in euphausiid and decapod larvae, provides the first evidence of a possible transformation of an apposition eye into a superposition eye during evolution. There are two different mechanisms by which superposition rays are formed (but not used) in the larval eyes (Fig. 2), and these mechanisms are retained and fully exploited by the adults (Fig. 1). Thus, euphausiids have refracting superposition eyes and decapods reflecting ones.

Recently Fincham⁴ and Land¹⁶ proposed that the different eye types in crustaceans are important indicators of taxonomic relationship; the present data indicate that this may be invalid for relations within the decapods. Not all adult decapods have superposition eyes (Table 1) but with the potential of the larval eyes to form superposition eyes, these may have evolved several times within the group without any correlation with taxonomic relationship. It can also be concluded that as a planktonic camouflage strategy is the probable origin of superposition eyes in decapods and euphausiids, presumed taxonomic relations should be based on larval rather than adult eye types.

I thank the staff of the Kristineberg Marine Biological Station and Mr Stellan Karlsson at the Blentarps Aquatic Laboratory, for supplying the animals. I also thank Professor Rolf Elofsson and Drs Dick Nässel and Eric Hallberg for critically reading the manuscript. I acknowledge financial assistance from the Swedish Natural Science Research Council (grant 2760-106).

Received 23 February; accepted 14 March 1983.

- Exner, S. *Die Physiologie der Facettierten Augen von Krebsen und Insekten* (Deuticke, Leipzig, 1891).
- Kunze, P. in *Handbk of Sensory Physiology* Vol. 7/6A (ed. Autrum, H.) 441-502 (Springer, Berlin, 1979).
- Land, M. F. in *Handbk of Sensory Physiology* Vol. 7/6B (ed. Autrum, H.) 471-592 (Springer, Berlin, 1981).
- Fincham, A. A. *Nature* **287**, 729-731 (1980).
- Land, M. F. *Nature* **287**, 681-686 (1980).
- Cleary, P., Deichel, G. & Kunze, P. *J. comp. Physiol.* **119**, 73-84 (1977).
- Vogt, K. Z. *Naturf.* **30c**, 691 (1975).
- Land, M. F. *Nature* **263**, 764-765 (1976).
- Vogt, K. *J. comp. Physiol.* **135**, 1-19 (1980).
- Nilsson, D.-E., Andersson, M., Hallberg, E. & McIntyre, P. *J. Microsc.* (in the press).
- Bryceson, K. P. *J. exp. Biol.* **90**, 347-350 (1981).

12. Meggitt, S. & Meyer-Rochow, V. B. in *The Compound Eye and Vision in Insects* (ed. Horridge, G. A.) 314–320 (Clarendon, Oxford, 1975).
13. Nilsson, D.-E. *J. comp. Physiol.* **147**, 339–349 (1982).
14. Nilsson, D.-E. & Nilsson, H. L. *J. exp. mar. Biol. Ecol.* **68** (1983).
15. Nilsson, D.-E. *J. exp. Zool.* **225**, 161–165 (1983).
16. Land, M. F. in *Sense Organs* (eds Laverack, M. S. & Cosens, D. J.) 31–48 (Blackie, Glasgow, 1981).
17. Land, M. F. *J. comp. Physiol.* **145**, 209–226 (1981).
18. Nilsson, D.-E., Odselius, R. & Elofsson, R. *Cell Tissue Res.* **230** (1983).

Binocular stereopsis in an insect

Samuel Rossel

Department of Zoology, University of Zürich, CH-8057 Zürich, Switzerland

Binocularity in insects is generally assumed to have the same function as in many vertebrates—the perception of depth. Evidence for this hypothesis stems from the observation that one-eyed dragonfly larvae, tiger beetles, praying mantids and water scorpions rarely catch prey^{1–5}, but no definitive evidence is available. Depth perception and catching behaviour depend on visual attention and visual behaviour and it is difficult to assess what is impaired when one eye is occluded⁶. A more promising approach to studying the importance of binocular disparity is one that does not interfere with normal binocular vision, and allows potential monocular depth cues to be controlled carefully. These criteria were met in the present study by the use of prismatic lenses placed in front of the compound eyes of the praying mantis, thus creating a conflict between binocular disparity and monocular cues. The results demonstrate that mantids do indeed rely on binocular triangulation when estimating the distance of prey, and thus provide the first unequivocal evidence for stereoscopic vision in an invertebrate.

Observations were made in the common Australian mantis *Tenodera australasiae* and the European species *Mantis religiosa*. Both mantids have a pronounced forward-looking fovea and a very large binocular field extending more than 70° in the frontal eye region (for *Tenodera* see refs 7, 8). Experi-

mental animals were fixed to a holder and centred in a white, homogeneous cardboard box (Fig. 1). Behaviour was recorded from above with a video camera. Living flies, usually the blowfly *Sarcophaga*, were provided as prey and fixed to a metal rod which could be moved by the motor of an x/y recorder. In addition, the motor was driven by a random-noise source so as to produce small, jerky movements in the lateral direction. Such movements are important to maintain the mantis's attention and to increase the strike frequency. At the beginning of each trial the fly was placed at the periphery of the arena, ~200 mm from the mantis's head. Usually the target was quickly detected by the insect and fixated by saccadic movements of the head. Once fixation had occurred, the target was slowly moved towards the mantis until the strike of the raptorial forelegs was released. Later the corresponding target distance was read from the monitor and related to the interocular distance in order to determine the binocular disparity (Fig. 2A).

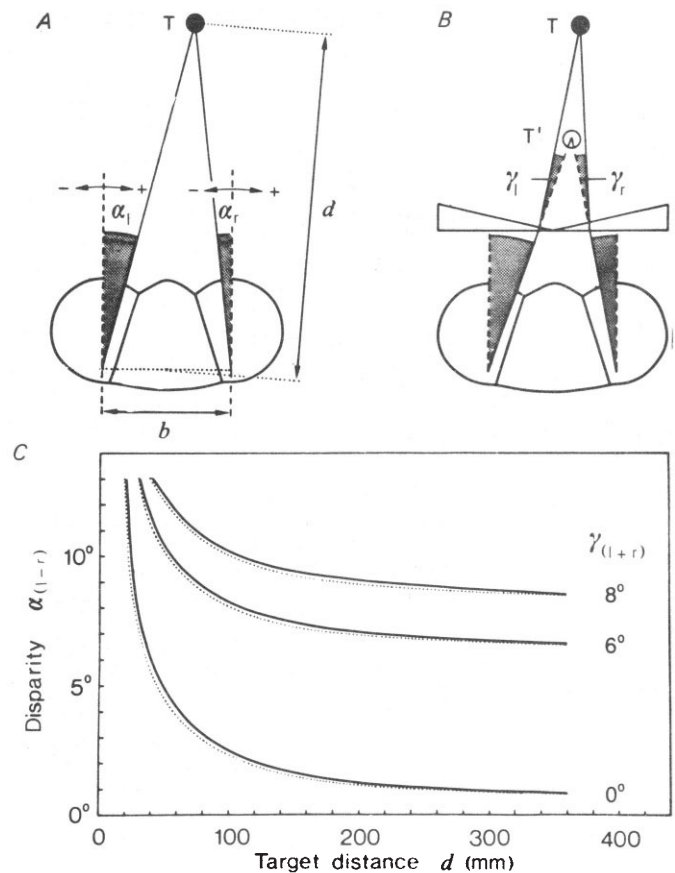


Fig. 2 A, Definitions. A praying mantis views a target (T) ahead. The angular positions of the corresponding retinal images (α_l for the left eye and α_r for the right eye) are referred to the forward-looking foveas. In considering the sign conventions, binocular disparity is then defined as the difference between α_l and α_r ($\alpha_{(l-r)}$). Accordingly, the binocular eccentricity of the target is given by the mean between α_l and α_r ($\alpha_{(l+r)/2}$). The interocular distance (b) corresponds to the horizontal separation between the two foveas. The target distance (d) is defined as the distance between the centre of the target and the point in the mantis head which lies between the optical centres of curvature of the two foveal eye regions. In the mantids studied here, these centres were 3.5 mm behind the eye surface, as calculated from facet diameters and interommatidial angles. B, Prisms inserted in front of the compound eyes deviate the lines of sight, thus leaving an insect which uses disparity as a cue to depth with an illusory target position (T'). γ_l and γ_r define the angular deviations produced by the left and right eye prism, respectively. The summed deviation power of both prisms is denoted by $\gamma_{(l+r)}$. C, Graph demonstrating the relationship between disparity and target distance for three different prism powers ($\gamma_{(l+r)} = 0^\circ, 6^\circ, 8^\circ$) and two different eccentricities ($\alpha_{(l+r)/2} = 0^\circ$, solid line; and $\alpha_{(l+r)/2} = 20^\circ$, dotted line). The interocular distance is 4.4 mm. This value is typical for the mantids used in this study.

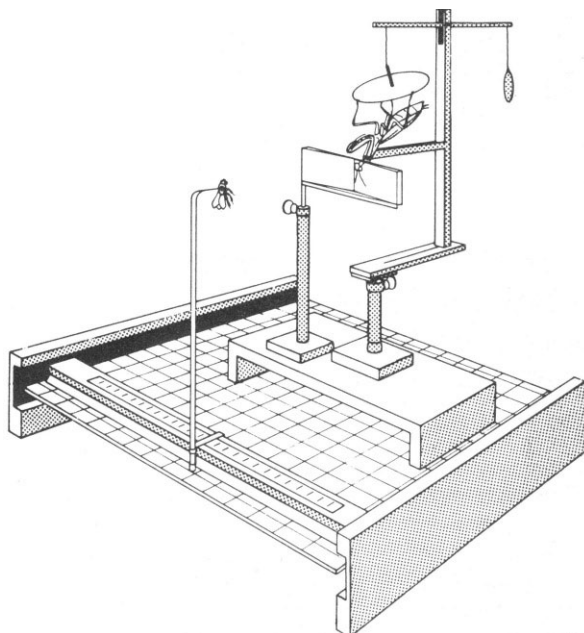


Fig. 1 A praying mantis views a fly through a pair of deviation prisms. The mantis is fixed upside down by the thorax; apart from this there is no restriction to body movements (the platform held by the legs can be rotated as well as moved up and down). Position and movement of the fly are controlled by an x/y plotter. During the experiments reported here, the speed with which the fly approached the mantis was gradually decreased with decreasing distance from the mantis head so as to keep the angular velocity of the retinal images as constant as possible (on the average the angular velocity was 0.3 deg s^{-1}).

Next, either one or two prisms, with their broad bases outward, were fixed on a separate holder and accurately aligned 2 mm in front of the forward-looking eyes (Figs 1, 2B). Obviously, this arrangement of prisms increases retinal disparity, without altering monocular depth cues such as motion parallax and image size (Fig. 2B, C). Thus, if the mantids are using binocular triangulation, they should strike short of the target by an amount that is positively correlated with the prism strength. The results demonstrate that there is indeed a correlation between prism strength and target distance at which strikes are released, such that the mantids strike when the disparity of the retinal images reaches some fixed value (Fig. 3A). Neither of the monocular cues to depth could explain the behaviour observed, and we conclude that mantids rely on binocular disparity when estimating the distance of prey.

In the experiments described above, the mantids could freely move the head and thus were able to fixate the target with their forward-looking foveas. Binocular depth perception, however,

is not restricted to the frontal field of view, but operates over most of the binocular field. This is demonstrated by using mantids in which the head is firmly glued to the thorax in order to prevent fixation of eccentric targets. The strike frequency then drops drastically with increasing eccentricity of prey; however, the disparity of the retinal images at which strikes are released does not change, clearly indicating that prey distance is still acquired by binocular triangulation (Fig. 3B).

It should also be noted that only the first strike released by the approaching target has been taken into account. Any further approach of prey evokes additional strikes, where a number of temporal and spatial characteristics of foreleg movement are correlated with the target distance⁹. This suggests that retinal disparity provides distance information over a much wider range than is expressed here. Limits are imposed, however, by the small interocular distance which allows good depth resolution only with respect to near targets (Fig. 2C and ref. 10).

Of course, it is possible that the praying mantis also uses other cues than disparity to estimate range. In fact, mantids often inspect stationary objects by performing extensive side-to-side body movements, which might be designed to extract motion parallax information, as has been suggested for the peering movements in other insects¹¹⁻¹³. Motion parallax, however, does not provide simple cues, when the object of interest is itself in motion. Moreover, in mantids neither convergence nor accommodation is available as a cue for the perception of depth. Thus, it appears that the absolute distance of rapidly moving prey can be assessed only by binocular disparity.

Binocular vision thus provides the praying mantis with a powerful means of locating objects in space. Stereopsis, therefore, is not an exclusive attribute of vertebrate vision, but has also evolved in invertebrates.

I thank H. U. Eugster for manufacturing the prisms, and R. Wehner and M. V. Srinivasan for reading the manuscript. This work was supported by Swiss NSF grant 3.073.0.81 (R. Wehner).

Received 6 December 1982; accepted 9 March 1983.

1. Baldus, K. Z. *vergl. Physiol.* **3**, 475-505 (1926).
2. Friederichs, H. F. Z. *Morph. Ökol. Tiere* **21**, 1-172 (1931).
3. Maldonado, H. & Levin, L. Z. *vergl. Physiol.* **56**, 258-267 (1967).
4. Maldonado, H. & Barros-Pita, J. C. Z. *vergl. Physiol.* **67**, 58-78 (1970).
5. Cloarec, A. *Biol. Behav.* **4**, 173-191 (1978).
6. Via, S. E. J. *comp. Physiol.* **121**, 33-51 (1977).
7. Rossel, S. J. *comp. Physiol.* **131**, 95-112 (1979).
8. Rossel, S. J. *comp. Physiol.* **139**, 307-331 (1980).
9. Corrette, B. J. thesis, Univ. Oregon (1980).
10. Burkhardt, D., Darnhofer-Demar, B. & Fischer, K. J. *comp. Physiol.* **87**, 165-188 (1973).
11. Wallace, G. K. J. *exp. Biol.* **36**, 512-525 (1959).
12. Horridge, G. A. *Endeavour* **1** (1), 7-17 (1977).
13. Collett, T. S. J. *exp. Biol.* **76**, 237-241 (1978).

Alteration of leukotriene D₄ hypotension by thyrotropin releasing hormone

Warren E. Lux Jr, Giora Feuerstein & Alan I. Faden

Neurobiology Research Unit, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814, USA

Leukotriene D₄ (LTD₄), a component of slow reacting substance of anaphylaxis (SRS-A)¹, produces marked hypotension in laboratory animals^{2,3}, implicating it as a potential mediator of anaphylactic shock. It has been demonstrated that naloxone reverses the hypotension associated with endotoxaemia⁴⁻⁶, hypovolaemia^{7,8} and spinal injury^{9,10}, presumably through blockade of endogenous opioid systems¹¹. More recently, it has also been shown that thyrotropin releasing hormone (TRH) improves experimental shock¹² and spinal injury¹³, and it has been postulated that TRH acts by 'physiological' antagonism of endogenous opioid systems in these conditions. Here we report that TRH both blocked and reversed leukotriene-induced hypotension in the unanaesthetized guinea pig, whereas

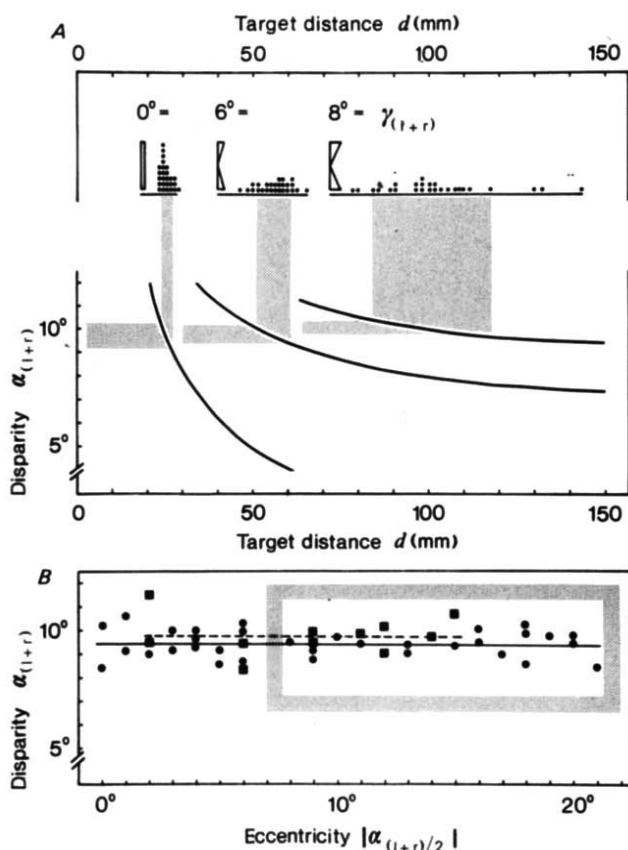


Fig. 3 A, The upper part of the graph shows the target distance at strike release for a *T. australasiae* whose vision has been distorted with base-out prisms of increasing strength ($\gamma_{(1+r)} = 0^\circ, 6^\circ, 8^\circ$). Each data point refers to a single strike. Following each strike the target was returned to its original position at the periphery of the arena and then again moved towards the mantis. In the lower part of the graph disparity is plotted against target distance for the prism powers used. The experimental data show that all strikes were released within the same small range of disparity regardless of prism strength. This indicates that distance information has been derived by binocular triangulation. The grey areas give the standard deviations of the target distance (vertical parts) and the disparity (horizontal parts) around the mean. B, The disparity at strike release as a function of the eccentricity of the target. The data were gathered from a *M. religiosa* specimen whose head was fixed relative to the thorax. The circles and the solid regression line represent results when no prisms were used. The squares and the broken regression line refer to data obtained when the right eye was occluded by a deviation prism ($\gamma_r = 6^\circ$). For the data points framed by the grey area the target was completely out of the median plane of the mantis head. As the disparity at which strikes are released is constant regardless of the eccentricity of the target, it can be concluded that binocular triangulation operates over a wide range of the binocular field.

naloxone had no effect. LTD₄ hypotension was also reversed by intracerebroventricular (i.c.v.) administration of TRH at a dose that had no effect when given systemically. LTD₄ administration was associated with sympatho-adrenomedullary activation, and TRH further augmented this response. Peripheral cholinergic blockade with methylatropine did not alter the leukotriene hypotension. These data demonstrate the first dissociation of TRH and naloxone in experimental shock and suggest that the beneficial effects of TRH in this model result from central interactions which are not mediated by endogenous opioids. The findings further provide a potential link between the physiological (that is, cardiovascular) effects of a peptide (TRH) and those of a leukotriene; this may have implications with regard to the pathophysiology and therapy of anaphylaxis.

Male guinea pigs (Hartley strain; 500–700 g) were anaesthetized with xylazine (4 mg per kg subcutaneous, s.c.) and ketamine (60 mg per kg intramuscular, i.m.). The femoral artery and vein were cannulated using PE-50 polyethylene tubing, and the cannulae were externalized by the method of Chiueh and Kopin¹⁴ for subsequent drug administration and cardiovascular monitoring. Animals to be used for i.c.v. injection studies also had indwelling intraventricular cannulae stereotactically placed at this time. All studies were performed 24 or more hours after surgery at which time the animals were fully awake and freely moving in their home cages. At the time of study, blood pressure was measured from the arterial cannula using a microtransducer (Narco Bio-Systems, model RP1500i) connected to a strain gauge coupler (Narco Bio-Systems, type 7179). Heart rate was obtained from the same signal using a biotachometer coupler (Narco Bio-Systems, type 7302). Blood pressure and heart rate were both continuously recorded and displayed on a physiograph (Narco Bio-Systems, model MK-IV). LTD₄ at 5 µg per kg intravenous (i.v.), a dose shown to produce shock in pilot studies, was given in a 0.2-ml bolus to all animals. In the pretreatment (blockade) studies, the control group received LTD₄ alone (seven animals), and the experimental groups received either TRH¹² (Beckman; 2 mg per kg i.v., *n* = 7), naloxone⁵ (Endo; 2 mg per kg i.v., *n* = 7) or *N*-methylatropine¹⁵ (Sigma; 2 mg per kg i.v., *n* = 6). In the post-treatment (reversal) studies, the animals received either TRH (Bachem; 1 mg per kg i.v., *n* = 6), naloxone (Endo; 2 mg per kg i.v., *n* = 6) or an equal volume of physiological saline (*n* = 6). In the i.c.v. reversal studies, animals received either TRH (Beckman; 20 µg per kg i.c.v., *n* = 6) or an equal volume of physiological saline (20 µl per kg i.c.v., *n* = 6). In the catecholamine studies, animals received either LTD₄ alone (six animals) or LTD₄ with TRH pretreatment (six animals), and arterial blood was obtained from each animal 30 s and 5 min after LTD₄, as well as during the baseline period before any drug administration. The samples were then processed and the plasma assayed for catecholamine levels by a radioenzymatic thin-layer chromatographic procedure described by DaPrada and Zürcher¹⁶ and by Weise and Kopin¹⁷.

LTD₄ (5 µg per kg i.v.) produced a biphasic blood pressure response in unanaesthetized guinea pigs; a transient period of

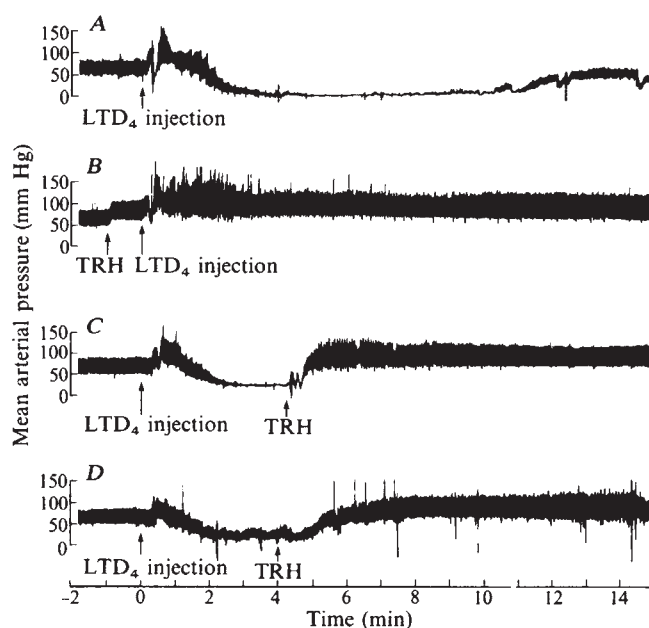


Fig. 1 Effect of TRH on LTD₄ hypotension in the unanaesthetized guinea pig. **A**, LTD₄ (5 µg per kg i.v.) produces profound hypotension in control animals. **B**, This hypotension is completely blocked by TRH (2 mg per kg i.v.). It is also reversed by TRH administered either **C**, systemically (1 mg per kg i.v.) or **D**, centrally (20 µg per kg i.c.v.). The figure shows an actual physiograph tracing from a representative animal in each group.

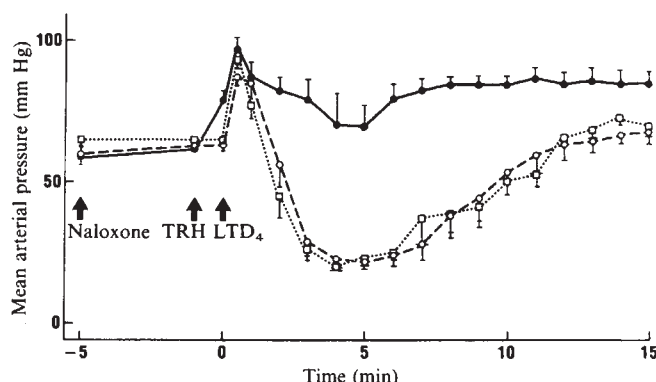


Fig. 2 Effect of pretreatment with TRH (2 mg per kg i.v.; ●) compared with naloxone (2 mg per kg i.v.; □) on the hypotension produced by LTD₄ (5 µg per kg i.v.). ○, Controls. TRH blocked the LTD₄-induced hypotension while naloxone had no effect. Results are the mean $\pm$ s.e.m. for seven animals in each group.

Table 1 Effect of TRH on increase in plasma catecholamine levels after LTD₄ administration

	Time after LTD ₄ injection	Noradrenaline (pg ml ⁻¹)	Adrenaline (pg ml ⁻¹)	Dopamine (pg ml ⁻¹)
Controls	30 s	281 $\pm$ 230	1,600 $\pm$ 798	22 $\pm$ 22
	5 min	670 $\pm$ 368	2,376 $\pm$ 861	110 $\pm$ 77
TRH pre-treatment	30 s	1,449 $\pm$ 738	1,069 $\pm$ 188	142 $\pm$ 57
	5 min	8,430 $\pm$ 2,850	6,016 $\pm$ 1,430	548 $\pm$ 129

Results are the mean $\pm$ s.e.m. for six animals in each group. Data represent increases over baseline values drawn before any pharmacological manipulation.

hypertension was followed by a profound and more prolonged period of hypotension lasting 10–15 min (Fig. 1A). TRH blocked the hypotensive but not the hypertensive phase of this response, whereas naloxone had no effect on either phase (Fig. 2). Differences between TRH-treated and either naloxone-treated or control animals were significant throughout the hypotensive phase (repeated measurement analysis of variance, ANOVA: $F = 15.08$, $P < 0.01$ for TRH compared with naloxone; $F = 10.99$, $P < 0.01$ for TRH compared with controls; each *n* = 7). TRH, administered i.v. at the nadir of the LTD₄ hypotension, also promptly reversed this hypotension, and naloxone continued to have no effect (Fig. 3A). Again, differences between TRH and either naloxone or control animals were significant (repeated measurement ANOVA: $F = 18.52$, $P < 0.01$ for TRH compared with naloxone; $F = 19.87$, $P < 0.01$ for TRH compared with controls; each *n* = 6). In addition, significant reversal of the LTD₄ depressor response

(repeated measurement ANOVA, $F = 17.89$, $P < 0.01$, $n = 6$) was produced when the TRH was given i.c.v. (Fig. 3B). The dose of TRH used in the i.c.v. studies had no effect when administered systemically. Moreover, i.c.v. administration of MK771 (a synthetic TRH analogue) to two animals at the same dose (20 µg per kg) produced reversal comparable to that seen with i.c.v. TRH.

Both the hypertensive and hypotensive phases of the LTD₄-induced blood pressure response were accompanied by bradycardia. TRH significantly blocked the bradycardia associated with the hypotensive phase (repeated measurement ANOVA: $F = 19.36$, $P < 0.01$ for TRH compared with naloxone; $F = 28.01$, $P < 0.001$ for TRH compared with controls; each $n = 7$), but not that associated with the hypertensive phase. Naloxone again had no effect. *N*-methylatropine had no effect on the blood pressure response to LTD₄, although it did block both phases of the bradycardia (data not shown).

Plasma catecholamines increased after LTD₄ administration during both the pressor and depressor phases (Table 1). In TRH-pretreated animals, some catecholamines were further increased at 30 s after LTD₄ and all were markedly elevated at 5 min after LTD₄. At 5 min, all three catecholamines measured were significantly elevated in TRH-treated compared with control animals (Student's *t*-test: $t = 2.72$, $P < 0.05$ for noradrenaline; $t = 2.55$, $P < 0.05$ for adrenaline; $t = 3.41$, $P < 0.01$ for dopamine; each $n = 6$).

The present data demonstrate for the first time a dissociation between the effects of TRH and naloxone on blood pressure in a model of experimental shock. The failure of naloxone to block or reverse the hypotensive response in this model suggests that the effect of TRH was unrelated to its ability to act *in vivo* as a 'physiological' opiate antagonist as postulated previously¹².

Moreover, the inability of methylatropine to block the hypotensive response indicates that the ability of TRH to block LTD₄ hypotension is unlikely to be mediated through a parasympathetic mechanism, as previously demonstrated for naloxone in spinal shock¹⁵. The fact that TRH enhanced LTD₄-induced catecholamine release is consistent with previous observations that TRH activates the sympatho-adrenomedullary system^{18,19}. However, this also seems unlikely to account for the present therapeutic effect of TRH as the pressor effects of TRH at the doses used in this study have previously been shown to be independent of associated catecholamine release¹⁹. Further support for this conclusion comes from the observation that indomethacin, which enhances LTD₄-induced catecholamine release to a greater degree than TRH, does not improve the hypotension in this model (in review).

The TRH effect does, however, seem to be mediated centrally, as LTD₄ hypotension was reversed by i.c.v. administration of TRH at a dose which had no effect on that hypotension when given systemically. Further evidence for a central site of action of TRH is provided by recent studies from this laboratory which demonstrate that LTD₄-induced hypotension in the SHR rat is blocked by TRH in the intact animal but not in pithed preparations in which central pathways are destroyed (Feuerstein, G., Zukowska, Z. & Faden, A. I., in preparation). The specific central site of TRH action in the present model is unknown, but TRH-like immunoreactivity has been demonstrated in several important brain cardioregulatory centres including the anterior hypothalamus, nucleus tractus solitarius and nucleus ambiguus^{20,21}. Moreover, the fact that the TRH effect can be reproduced by a synthetic TRH analogue (MK771) suggests further that a central receptor-mediated event may be involved.

LTD₄-induced hypotension in the unanaesthetized guinea pig has been unresponsive to treatment with cyclooxygenase inhibitors and antioxidants (in review), therapies which have proven efficacious in other types of experimental shock^{22,23}. Similarly, naloxone proved totally ineffective in the present model. Thus, the ability of TRH to block LTD₄ hypotension seems to be relatively unique and suggests a potential non-endorphin-mediated link between the cardiovascular effects of this leukotriene and those of a centrally acting peptide (TRH). Given the LTD₄-SRS-A association, moreover, such a link may ultimately have implications for the development of new therapeutic approaches to anaphylactic shock.

We thank Mr George Smith and CPT Fair Meeks for technical assistance, and Miss Eleanor M. Bell for preparation of the manuscript. This work was supported by the Uniformed Services University of the Health Sciences Protocol no. R09201. The experiments reported here were done according to the principles set forth in National Research Council DHEW publication no. 74-23.

Received 17 January; accepted 28 February 1983.

- Lewis, R. A. & Austen, K. F. *Nature* **293**, 103–108 (1981).
- Drazen et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4354–4358 (1980).
- Feuerstein, G., Zukowska-Grojec, Z. & Kopin, I. J. *Eur. J. Pharmac.* **76**, 107–110 (1981).
- Holaday, J. W. & Faden, A. I. *Nature* **275**, 450–451 (1978).
- Faden, A. I. & Holaday, J. W. *J. Pharmac. exp. Ther.* **212**, 441–447 (1980).
- Reynolds, D. G. et al. *Circulatory Shock* **7**, 39–48 (1980).
- Faden, A. I. & Holaday, J. W. *Science* **205**, 317–318 (1979).
- Vargish, T. et al. *Circulatory Shock* **7**, 31–38 (1980).
- Holaday, J. W. & Faden, A. I. *Brain Res.* **189**, 295–299 (1980).
- Faden, A. I., Jacobs, T. P. & Holaday, J. W. *Science* **211**, 493–494 (1981).
- Faden, A. I. & Holaday, J. W. *J. infect. Dis.* **142**, 229–238 (1980).
- Holaday, J. W., D'Amato, R. J. & Faden, A. I. *Science* **213**, 216–218 (1981).
- Faden, A. I., Jacobs, T. P. & Holaday, J. W. *New Engl. J. Med.* **305**, 1063–1067 (1981).
- Chiuah, C. C. & Kopin, I. J. *Am. J. Physiol.* **234**, H690–H695 (1978).
- Faden, A. I., Jacobs, T. P. & Holaday, J. W. *J. Autonomic Nervous Syst.* **2**, 295–304 (1980).
- DaPrada, M. & Zurcher, G. *Life Sci.* **19**, 1161–1174 (1976).
- Weise, V. K. & Kopin, I. J. *Life Sci.* **19**, 1673–1686 (1976).
- Yarabrough, G. C. in *Prog. Neurobiol.* Vol. 12 (eds Kerkut, G. A. & Phillis, J. W.) 291–312 (Pergamon, Oxford, 1979).
- Horita, A., Carino, M. A. & Weitzman, R. E. in *Catecholamines: Basic and Clinical Frontiers* Vol. 2 (eds Usdin, E., Kopin, I. J. & Barchas, J.) 1140–1142 (Pergamon, New York, 1979).
- Rea, M. A., Kubek, M. J., Hodes, Z. I. & Aprison, M. H. *Soc. Neurosci. Abstr.* **8**, 110 (1982).
- Lechan, R. M. & Jackson, I. M. D. *Soc. Neurosci. Abstr.* **8**, 111 (1982).
- Fletcher, J. R. & Ramwell, P. W. *Br. J. Pharmac.* **64**, 185–191 (1978).
- Lefer, A. M., Araki, H. & Okamoto, S. *Circulatory Shock* **8**, 273–282 (1981).

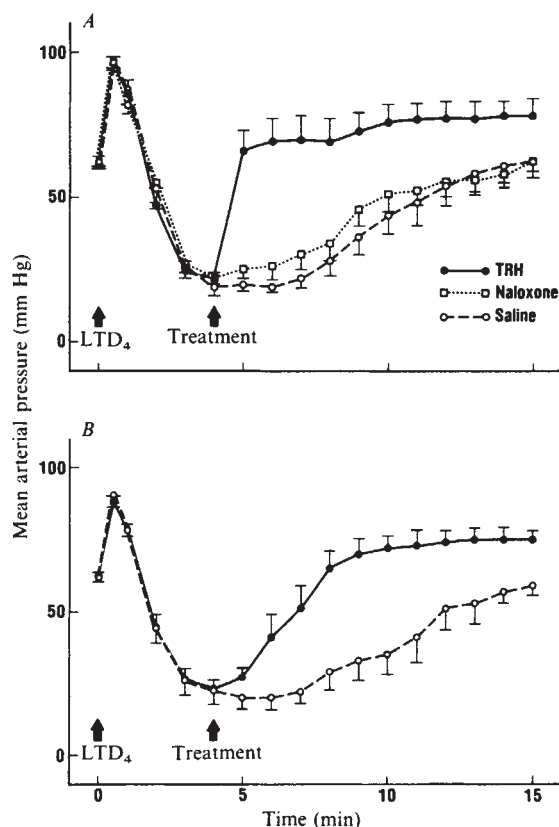


Fig. 3 Reversal of LTD₄ hypotension by TRH (●) given either **A**, systemically (1 mg per kg i.v.) or **B**, centrally (20 µg per kg i.c.v.). ○, Controls. Systemic administration of naloxone at 2 mg per kg i.v. (□) had no effect (**A**). Results are the mean $\pm$ s.e.m. for six animals in each group.

Lymphokine-induced IgM secretion by clones of neoplastic B cells

Kathryn Brooks*, Dorothy Yuan*, Jonathan W. Uhr*, Peter H. Krammer† & Ellen S. Vitetta*

* Department of Microbiology, University of Texas Health Science Center, Dallas, Texas 75235, USA and † the Institute for Immunology and Genetics, German Cancer Research Center, Heidelberg, FRG

The induction of antibody secretion by B cells requires T-cell-derived factors¹⁻⁵. Such factors have been described^{1,2,6-12} but the precise relationship among these various factors is not clear, and it has been difficult to demonstrate that these factors act directly on the B cell and do not exert their effect via T cells or macrophages. In this report we describe the direct induction of IgM synthesis and secretion in cloned lines of long-term tissue culture adapted neoplastic B cells (BCL₁) by T-cell supernatants from phorbol-12-myristate 13-acetate (PMA)-induced EL-4 cells or concanavalin A (Con A)-induced 7.1.1a cells^{5,9}. We have termed this activity BCDF μ (B-cell differentiation factor for IgM). The supernatants containing BCDF μ induce activated and neoplastic B cells to secrete IgM⁵ and the factor responsible is distinct from BCGF¹³, interleukin-2 (IL-2)⁵, the classical T-cell replacing factor (TRF) described by Schimpl and Wecker⁵, and immune interferon (IFN γ)⁵.

A report by Gronowicz *et al.*¹⁴ indicated that the lipopolysaccharide (LPS)-inducible *in vitro*-adapted neoplastic B cells, BCL₁, did not generate plaque-forming cells (PFC) before treatment with LPS but in a more sensitive radioimmunoassay we found that considerable amounts of IgM were secreted by these cells even in the absence of LPS ($4.8 \pm 2 \mu\text{g ml}^{-1}$ after 4 days of culture at 5×10^5 cells per ml). Therefore, the five BCL₁ clones obtained from the *in vitro* cell line were selected for

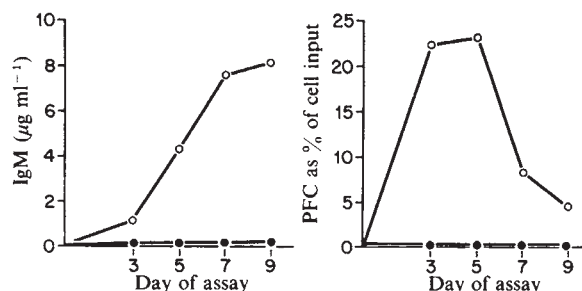


Fig. 1 Kinetics of IgM secretion induced by EL-4 SN as determined by both radioimmunoassay and PFC assay. 2×10^4 CW13.14-1G3 cells were cultured with (○) and without (●) an optimal concentration of EL-4 SN. The concentration of IgM in the culture SN and number of PFC were determined on days 3, 5, 7 and 9 after addition of the T-cell supernatant. A reverse plaque assay²⁶, using protein A coupled to sheep erythrocytes followed by incubation with RαM μ and guinea pig complement was used for the PFC determination. See legend to Table 2 for experimental details.

their lack of spontaneous IgM secretion ($<0.5 \mu\text{g ml}^{-1}$). Two initial clones were subcloned and the probability that the sub-clones were derived from a single BCL₁ B cell was $>97\%$ for clones 1F8, 1G3 and 4D7 and $>99.9\%$ for clones 3B3 and 3G2.

The cell surface phenotype of the five BCL₁ clones is similar (Table 1). They express high levels of surface IgM with the BCL₁ idiotype. A portion of the cells express Ia antigens (I-A on 30–70% and I-E antigens on 20–50% of the cells) and surface (s) IgD (15–25% of the cells). The variability in Ia and IgD antigen expression may be due to cell cycle-related fluctuations in antigen expression as shown for Ia antigens on the B-cell line, WEHI-231¹⁵. The cell surface phenotype of these BCL₁ clones is identical to that of the *in vivo* BCL₁ cells¹⁶.

All five clones of BCL₁ cells differentiate into antibody-forming cells when cultured with supernatants from either the PMA-induced EL-4 T lymphoma cells or the Con A-induced alloreactive (AKR anti-B6), PK7.1.1a T cell line. Six days after addition of these supernatants, the concentration of IgM in the medium had increased 15–80-fold, reaching maximum levels of $13 \mu\text{g ml}^{-1}$ (Table 2). At the peak of this response, 70–80% of the cells were viable and a twofold increase in the cell number had occurred. The addition of LPS or an IL-1 enriched supernatant from LPS-stimulated PU5-1.8 cells did not enhance the response (data not shown). Maximal PFFFC were detected 5 days after the addition of EL-4 supernatant and were equivalent to 20–25% of the input cells (Fig. 1).

The level of IgM secretion induced by supernatant from PK7.1.1a was lower than that induced by supernatant from EL-4. As this might have been due to the presence of residual PMA in the EL-4 supernatant we prepared a supernatant which consisted of the 50–85% saturated $(\text{NH}_4)_2\text{SO}_4$ (SAS) fraction of the EL-4 induction media (minus the EL-4 cells). This supernatant should contain at least as much PMA as in the PMA-induced EL-4 supernatant. However this supernatant was not only inactive when used alone but it had no effect on the amount of IgM secreted in response to PK7.1.1a supernatant (data not shown). Thus, residual PMA does not appear to be responsible for the differential response of the BCL₁ clones to EL-4 and PK7.1.1a supernatants.

To determine whether we were monitoring the same biological activity in both the EL-4 and PK7.1.1a supernatants, BCDF μ was partially purified by gel filtration¹³. The BCDF μ activity from the columns was determined by its ability to induce IgM secretion in the *in vivo* BCL₁ cells⁵. In both the EL-4 and 7.1 supernatants, the BCDF μ activity was found in a broad peak of 30–60,000 (30–60K) molecular weight (MW). When these partially purified BCDF μ preparations were assayed on the BCL₁ clones, they induced equivalent amounts of IgM secretion. However, the level of IgM secretion induced was about half of that observed using unfractionated EL-4 supernatant. The activity induced by unfractionated PK7.1.1a super-

Table 1 Surface phenotype of BCL₁ clones

Clone	BCL ₁ Id*	% Positive cells for			
CW13		sIgM†	sIgD‡	I-A§	I-E/C§
14-1F8	99	91	16	70	41
14-1G3	99	91	25	32	41
14-4D7	98	97	20	50	53
20-3B3	99	94	15	63	18
20-3G2	98	97	25	43	28

The BCL₁ tumour line described by Gronowicz *et al.*¹⁴ was cloned at limiting dilution in 100 μl cultures using 96-well microtitre plates (Flow). The 10% fetal calf serum (FCS)-RPMI 1640 media used for cloning was supplemented with L-glutamine, sodium pyruvate, non-essential amino acids, gentamycin, antibiotic-antimycotic solution (Gibco) as well as 25% v/v of a 4-day supernatant from the WEHI 274 tumour line. The cultures were incubated at 37 °C in 5% CO₂ and fed on day 10 with 100 μl of 10% FCS-RPMI media lacking WEHI 274 supernatants. BCL₁ clones which did not spontaneously secrete IgM ($<0.5 \mu\text{g ml}^{-1}$ after 5 days) were selected for further investigation. $1-2 \times 10^6$ BCL₁ cells were incubated for 30 min at 4 °C with either rabbit anti-BCL₁ idiotype (RαBCL₁ Id), affinity purified rabbit anti-mouse μ (RAM μ), anti-mouse δ (hybridoma H10.4.22), anti-Ia.7 (hybridoma 13/18), or anti-I-A (hybridoma 17/227). After washing in phosphate-buffered saline-azide, a secondary fluorescein isothiocyanate-coupled (F1) goat anti-rabbit immunoglobulin was added to RAM μ and RαBCL₁ Id-treated cells while F1-F(ab')₂ rabbit anti-mouse γ was used with the other primary antibodies. After washing and fixing the cells with 4% buffered paraformaldehyde, they were analysed on the fluorescence activated cell sorter (Becton-Dickinson FACS III)²⁵. Each point represents one to three determinations.

* Normal rabbit immunoglobulin gave a background of $2.5 \pm 0.6\%$ positive cells.

† s.e.m. $\leq 8\%$.

‡ Normal mouse serum gave a background of $9 \pm 1.4\%$ positive cells.

§ s.e.m. $\leq 23\%$.

Table 2 Induction of IgM secretion

CW13 clone	Treatment	No. of experiments	Stimulation index (mean $\pm$ s.e.)
14-1F8	LPS	2	3.4 $\pm$ 1.7
	EL-4 SN	5	84 $\pm$ 46
	PK 7.1 SN	2	13 $\pm$ 1.4
14-1G3	LPS	3	2.8 $\pm$ 1.4
	EL-4 SN	7	40 $\pm$ 20
	PK 7.1 SN	4	16 $\pm$ 12
14-4D7	LPS	3	2 $\pm$ 0.4
	EL-4 SN	7	14 $\pm$ 5
	PK 7.1 SN	3	9 $\pm$ 4
20-3B3	LPS	3	3.1 $\pm$ 1.1
	EL-4 SN	6	17 $\pm$ 8
	PK 7.1 SN	3	12 $\pm$ 7
20-3G2	LPS	3	4.4 $\pm$ 1.5
	EL-4 SN	8	17 $\pm$ 5
	PK 7.1 SN	3	17 $\pm$ 11

LPS was used at $20 \mu\text{g ml}^{-1}$. The preparation of both the EL-4 supernatant (SN) and the PK 7.1 SN have been described previously⁹. Briefly, after a 48 h induction with PMA the 85% SAS fraction of EL-4 SN was dissolved in 10% of its original volume and was dialysed against RPMI 1640. The alloreactive PK 7.1 cells were pulsed with Con A for 3 h, washed and SN was collected 48 h later. Cloned BCL₁ cells were plated at $2-5 \times 10^4$ cells per well in 96-well microtitre plates in 200 μl of 3-5% FCS-RPMI 1640 media. The cells were usually maintained in 3-5% FCS RPMI media for three to five generations before the addition of T-cell SN. They were stimulated with either $20 \mu\text{g ml}^{-1}$ LPS or an optimal dose of SN from EL-4 or PK 7.1. IgM secretion was assessed by a radioimmunoassay using rabbit anti-mouse immunoglobulin (RAMiG) coated plates and ^{125}I -R $\alpha\text{M}\mu$ ⁹. The stimulation index is the IgM concentration in SN of induced cells divided by IgM concentration in SN of uninduced cells. The standard error represents the variation between experiments.

nanant is also about half of that induced by EL-4 supernatant. This suggests that the EL-4 SN contains an additional factor (not residing in the 30-60K MW range) capable of augmenting the response induced with 30-60K MW material. The nature of this additional factor is unclear, but it is probably not IL-2¹⁷.

Before the addition of T-cell supernatant from EL-4, mRNA for membrane μ chain was the predominant species of cytoplasmic μ -specific RNA in all of the BL clones. After addition, d increase in the quantity of cytoplasmic mRNA encoding secreted μ chains (Fig. 2). Therefore, these data suggest that EL-4 supernatant can induce differentiation, rather than enhance the release of 'preformed' antibody from already differentiated cells¹⁸. Similar experiments were not performed with 7.1 supernatant for technical reasons, although such experiments are in progress.

The state of activation of the normal B-cell targets for BCDF μ -containing supernatant is unknown. Most probably, resting B cells must be activated by cross-linking of surface immunoglobulin receptors before they can respond to BCDF μ ^{3,19-21}. In addition, it has been suggested that BCGF²², and IL-1^{23,24} play a role in B-cell activation. In this regard, it should be noted that during cloning, the BCL₁ cells used in this study were exposed to supernatant from a macrophage-like tumour line, WEHI-274. This supernatant contains a low level of IL-1 (unpublished observations). Therefore, although IL-1 was not present during the induction period (supernatant from both EL-4 and PK 7.1 lack IL-1), the cells may have received an IL-1-mediated signal before treatment with BCDF μ -containing supernatant. As both the EL-4 and 7.1 supernatants contain BCGF¹³, it is not possible to determine whether BCGF plays a role in BCDF μ -mediated induction of IgM secretion. The cloned BCL₁ cells may represent analogues of B cells which have already received an initial activation signal, as they are neoplastic and are actively proliferating.

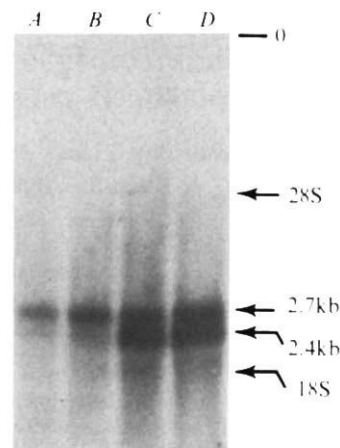


Fig. 2 Increase in cytoplasmic mRNA for μ chain of the secreted form of IgM after induction with EL-4 supernatant. RNA was prepared by hot phenol extraction from 5×10^6 cells which had been cultured in the absence (lanes A, B) or presence (lanes C, D) of EL-4 SN, size fractionated by agarose gel electrophoresis, transferred to nitrocellulose filter paper and hybridized with μ chain specific ^{32}P -labelled DNA probe (10^8 c.p.m. per μg) as previously described²⁷. Filters were washed and exposed to X-ray film at -70°C for 2 days. Molecular sizes were determined from the positions of 28S (4.7 kilobases, kb) and 18S (2.0 kb) ribosomal RNA visualized by acridine orange staining. Lane A contains RNA from clone 4D7. Lanes B, C are clone 1G3 and lane D is clone 3B3.

In conclusion, our data clearly demonstrate that supernatants derived from malignant and non-malignant T-cell lines act directly on malignant B cells to induce differentiation. These supernatants work in the absence of IL-2, conventional TRF, macrophages and T cells. These findings provide evidence that the target cells for BCDF μ -containing supernatants are B cells.

We thank Ms K. Sill, P. May, Mr P. Genthner, W. Muller and T. Tucker for technical assistance, Dr P. Tucker for the μ -specific cDNA plasmid probe, Dr Ellen Pure for helpful discussions, and Ms G. A. Cheek for secretarial assistance. Requests for the BCL₁ clones used in this study should be addressed to E.S.V. This work was supported by NIH grants AI-11851, AI-12789 and CA-31534.

Received 6 October 1982; accepted 16 February 1983.

- Dutton, R. W. *Transplant Rev.* **23**, 66-77 (1975).
- Schimpl, A. & Wecker, E. *Transplant Rev.* **23**, 176-201 (1975).
- Kishimoto, T., Miyake, T., Nishizawa, Y., Watanabe, T. & Yamamura, Y. *J. Immun.* **115**, 1179-1184 (1975).
- Parker, D. C., Fothergill, J. J. & Wadsworth, D. C. *J. Immun.* **123**, 931-941 (1979).
- Pure, E. *et al. J. Immun.* **127**, 1953-1958 (1981).
- Parker, D. C. *J. Immun.* **129**, 469-474 (1982).
- Armerding, D., Sachs, D. H. & Katz, D. H. *J. exp. Med.* **140**, 1717-1722 (1974).
- Delovitch, T. L., Biggin, J. & Fung, F.-Y. *J. exp. Med.* **147**, 1198-1212 (1978).
- Isakson, P. C., Pure, E., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **155**, 734-748 (1982).
- Liebson, H. J., Marrack, P. & Kappler, J. J. *J. Immun.* **129**, 1309-1402 (1982).
- Andersson, J., Schreier, M. H. & Melchers, F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1612-1616 (1980).
- Lernhardt, W., Corbel, C., Wall, R. & Melchers, F. *Nature* **300**, 355-357 (1982).
- Pure, E. *et al. J. exp. Med.* (in the press).
- Gronowicz, E. S., Doss, C. A., Howard, F. D., Morrison, D. C. & Strober, S. *J. Immun.* **125**, 976-980 (1980).
- Lanier, L. L. & Warner, N. L. *J. Immun.* **126**, 626-631 (1981).
- Krolick, K. A., Isakson, P. C., Uhr, J. W. & Vitetta, E. S. *Immun. Rev.* **48**, 81-106 (1979).
- Pure, E. *et al. J. Immun.* **129**, 2420-2425 (1982).
- Lafrenz, D., Koretz, S., Stratte, P. T., Ward, R. B. & Strober, S. *J. Immun.* **129**, 1329-1335 (1982).
- Parker, D. C. *Nature* **258**, 361-363 (1975).
- Yoshizaki, K. *et al. J. Immun.* **128**, 1296-1301 (1982).
- Julius, M. H., von Boehmer, H. & Sidman, C. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1989-1993 (1982).
- Howard, M. *et al. J. exp. Med.* **155**, 914-923 (1982).
- Hoffman, M. K. *J. Immun.* **125**, 2076-2081 (1980).
- Howard, M. & Paul, W. E. *A. Rev. Immun.* **1** (in the press).
- Loken, M. B. & Herzenberg, L. A. *Ann. N.Y. Acad. Sci.* **254**, 163-174 (1975).
- Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 588-590 (1976).
- Yuan, D. & Tucker, P. W. *J. exp. Med.* **56**, 962-974 (1982).

Voltage and Ca^{2+} -activated K^+ channel in baso-lateral acinar cell membranes of mammalian salivary glands

Y. Maruyama, D. V. Gallacher & O. H. Petersen*

The Physiological Laboratory, University of Liverpool,
Brownlow Hill, PO Box 147, Liverpool L69 3BX, UK

Nervous or hormonal stimulation of many exocrine glands evokes release of cellular K^+ (ref. 1), as originally demonstrated in mammalian salivary glands^{2,3}, and is associated with a marked increase in membrane conductance^{1,4,5}. We now demonstrate directly, by using the patch-clamp technique⁶, the existence of a K^+ channel with a large conductance localized in the baso-lateral plasma membranes of mouse and rat salivary gland acinar cells. The K^+ channel has a conductance of ~ 250 pS in the presence of high K^+ solutions on both sides of the membrane. Although mammalian exocrine glands are believed not to possess voltage-activated channels^{1,7}, the probability of opening the salivary gland K^+ channel was increased by membrane depolarization. The frequency of channel opening, particularly at higher membrane potentials, was increased markedly by elevating the internal ionized Ca^{2+} concentration, as previously shown for high-conductance K^+ channels from cells of neural origin⁸⁻¹⁰. The Ca^{2+} and voltage-activated K^+ channel explains the marked cellular K^+ release that is characteristically observed when salivary glands are stimulated to secrete.

We used collagenase-isolated clusters of acinar cells (typically containing 10–20 cells) from mouse parotid and mandibular (submaxillary) glands as well as from rat parotid. Single-channel current recordings from the baso-lateral membranes were made *in situ* or in excised membrane patches as described previously^{6,11,12}. Good seals between pipettes and acinar surface membranes (>10 G Ω) were obtained in all the three glands studied. The patches were mostly silent when the pipette potential was held at 0 mV, but in some cases occasional short-lasting outward current steps were observed. The frequency of such events markedly increased or silent patches became active when the pipette potential was made negative. The effect of changing the membrane potential on the pattern of channel opening and closure was immediate, viewed on the time scale used in Figs 1, 3 and 4. Figure 1D shows short-lasting unitary outward current steps from a membrane patch depolarized by 20 mV *in situ*. The number of outward current steps markedly increased after further depolarization (Fig. 1C) and the duration of the channel openings increased (Fig. 1B, C). When the patch was depolarized by 60 mV the channel was almost constantly open (Fig. 1A). Such voltage-activated large outward current channels were observed in six patches from mouse parotid acinar cells, two patches from rat parotid acinar cells and eight from mouse mandibular acinar cells. There seemed to be no differences in the characteristics of the channels in the three glands. The increase in the probability of channel opening and the prolongation of open times associated with membrane depolarization were not due to Ca^{2+} influx through voltage-activated Ca^{2+} channels, as the same pattern was observed in experiments in which the pipette solution contained no Ca^{2+} (nominally Ca^{2+} free solution with 1 mM EGTA). The relationship between size of unitary current step (I) and the pipette potential (V_p) is shown in Fig. 2 (open circles). The membrane potential of these cells is ~ -50 to -70 mV (refs 4,13), that is, the apparent reversal potential corresponds to a membrane potential of -70 to -90 mV which is close to the K^+ equilibrium potential, but much more negative than the Cl equilibrium potential¹.

After excision of the isolated patch, using a normal extracellular solution in the bath, the large voltage-dependent channel was not apparent. In this condition, in which the main cation on both sides was Na^+ , unitary activity (Fig. 1E) resembled that previously observed in pancreatic acinar cells^{11,12}. The pattern of opening and closing was independent of the membrane potential. The I - V relation is shown in Fig. 2 (triangles) and corresponds to a single-channel conductance of ~ 30 – 35 pS, a value similar to that found for the Ca^{2+} -activated inward current channel in the pancreatic acinar plasma membranes¹¹. This channel was observed in two membrane patches from the mouse parotid and three patches from the mouse mandibular gland (all excised, inside-out with symmetrical Na^+ solutions).

Figure 3 shows single-channel current recordings from an excised inside-out membrane patch exposed to extracellular solution (NaCl, high Ca^{2+}) on the outside (pipette) and intracellular solution (KCl, Ca^{2+} -free) on the inside (bath). When the pipette potential was held at +40 mV (Fig. 3E), occasional short-lasting outward current steps were observed. Reducing the pipette potential (membrane depolarization) markedly

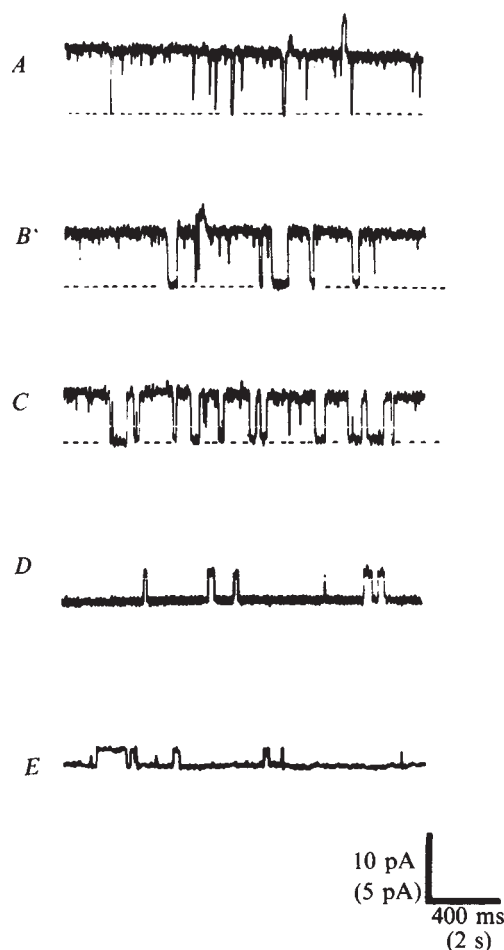


Fig. 1 Single-channel current traces from a baso-lateral membrane patch in collagenase-isolated mouse parotid acinar cell cluster. A–D were recorded *in situ* (cell-attached conformation) whereas E was from the same patch after excision (inside-out). The pipette was filled with standard extracellular solution having the following composition (mM): 140 NaCl, 4.7 KCl, 1.2 CaCl_2 , 1.13 MgCl_2 , 10 glucose and 10 HEPES. The solution was titrated to pH 7.2 with NaOH. The same solution was used in the bath. The patch resistance was 20 G Ω . The pipette potential was -60 mV in A, -50 mV in B, -40 mV in C and -20 mV in D. 150 Hz filtering (low pass) was applied. The horizontal broken lines represent the current level when all channels were closed. E, Record from same membrane patch after excision into the inside-out conformation⁶. Symmetrical extracellular solutions were used on both sides of the membrane. Pipette potential, -40 mV; low pass filtering at 30 Hz. Note that the calibration in E (given in parentheses) differs from that in A–D.

* To whom reprint requests should be addressed.

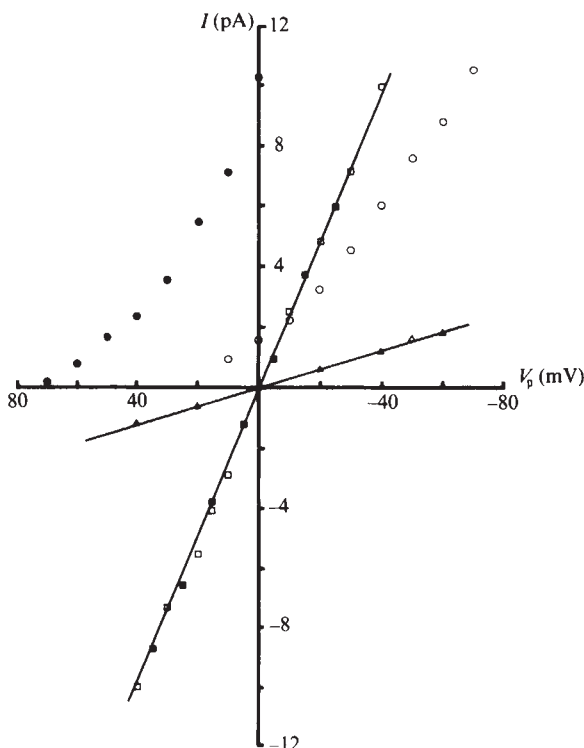


Fig. 2 Single-channel I - V relationships in different experimental situations. $\circ$, data from the *in situ* membrane patch (pipette filled with extracellular solution) represented by traces A-D in Fig. 1. Δ , Data from same patch after excision (inside-out) (symmetrical extracellular solutions on both sides of the membrane). $\bullet$, Data from excised inside-out patch exposed to extracellular solution on the outside (pipette) and to intracellular solution on the inside (bath). The intracellular solution had the following composition (mM): 145 KCl, 10 NaCl, 1.13 MgCl₂, 10 glucose, 10 HEPES; no calcium was added (but the solution contained $\sim 10 \mu\text{M Ca}^{2+}$) and EGTA was added (1 mM); pH 7.2. These data are from the mouse parotid membrane patch represented by the traces shown in Fig. 3. $\square$, Data from *in situ* membrane patch of K^+ -depolarized mouse mandibular acinar cell. The bath solution and pipette filling solution were similar to the intracellular solution except that there was no EGTA and the Ca^{2+} concentration was 1.2 mM. These data points are from the patch represented by the traces in Fig. 4A. $\blacksquare$, Data from the same patch after excision (inside-out). The bath solution (in contact with the intracellular aspect of the membrane) was the intracellular solution, that is, the same solution as in the pipette, but with 1 mM EGTA and no added Ca^{2+} . These data points correspond to the traces shown in (2) of Fig. 4C (only the large unitary steps were taken into account). V_p , Pipette potential.

increased the frequency of openings and their durations and at a membrane potential of 0 mV two channels were simultaneously open most of the time (Fig. 3A-D). The I - V relationship is shown in Fig. 2 (closed circles). The apparent reversal potential was -70 to -90 mV, which is close to the K^+ equilibrium potential. There was no Cl^- gradient across the membrane; the channel is therefore clearly K^+ -selective.

Figure 4A shows traces from an isolated patch *in situ*. High K^+ solution was present on the extracellular side of the patch (pipette) and also in the bath. The acinar cell from which the recording was made therefore had a membrane potential of ~ 0 mV (ref. 13). As there was no K^+ gradient, unitary current steps were not observed when the pipette potential was held at 0 mV (Fig. 4A, b). Membrane depolarization immediately evoked vigorous activity (Fig. 4A, a). Hyperpolarizing the patch also allowed a gradient to be established and thus unitary activity could be observed. However, with increasing hyperpolarization, the probability of opening and duration of open times declined (Fig. 4A, c-g). Figure 4B shows further traces from the same patch first *in situ* at a pipette potential of +15 mV

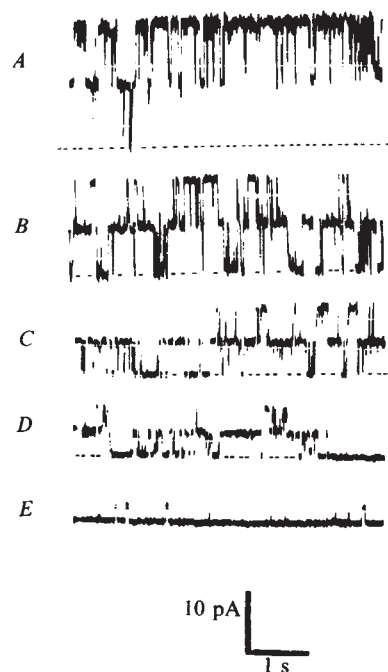


Fig. 3 Single-channel traces from an excised (inside-out) mouse parotid acinar membrane patch exposed to extracellular solution on the outside (pipette) and intracellular solution on the inside (bath). A-E show consecutive traces obtained at different pipette potentials: A, 0 mV; B, +10 mV; C, +20 mV; D, +30 mV; and E, +40 mV. Patch resistance 15 G Ω , low pass filtering at 150 Hz.

(a), then after excision (b-f) at the same pipette potential, but with different Ca^{2+} concentrations on the inside of the membrane (bath). Changing the bath Ca^{2+} concentration had an effect on the pattern of channel opening and closure within 5-10 s. Increasing the 'internal' (bath) ionized Ca^{2+} concentration from $<10^{-9}\text{M}$ to $\sim 10^{-7}\text{M}$ markedly increased the frequency of channel openings and their durations and at $3 \times 10^{-6}\text{M}$ the channels were almost constantly open. Smaller unitary steps were also encountered. Such additional small steps were often observed in patches exposed to high K^+ solution on both sides (see also Fig. 4C). These small steps probably represent a small K^+ channel as they were not observed in symmetrical Na^+ solutions. The conductance of this channel is higher (~ 50 pS) than that of the 30-35 pS channel observed in symmetrical Na^+ solutions (Fig. 1E). The I - V plots for K^+ channels shown in Fig. 2 relate exclusively to the large K^+ channel. Figure 4C shows traces from the same patch still exposed to high K^+ solutions on both sides of the membrane, comparing the effects of two different internal Ca^{2+} concentrations— 10^{-7}M in (1) and $<10^{-9}\text{M}$ in (2)—on the voltage sensitivity of the channel. The channels were more sensitive to voltage changes at the very low internal Ca^{2+} concentration. The Ca^{2+} activation of the K^+ channel was particularly dramatic when the membrane was slightly hyperpolarized (for example, compare (1) and (2) of Fig. 4C, c). The I - V relationships for situations with high K^+ concentration on both sides of the membrane (*in situ* and excised) are shown in Fig. 2 (squares). The two lines are identical, the slope corresponding to a single-channel conductance of ~ 250 pS.

Evidence from microelectrode recording studies in many different mammalian exocrine glands including the salivary glands, suggests the absence of voltage-activated channels^{1,7,14}. In the pancreatic acinar cell (mouse and rat) single-channel current recording studies have confirmed this hypothesis and so far only one type of channel, the voltage-insensitive Ca^{2+} activated non-discriminatory cation channel (30-35 pS)^{11,15,16} has been demonstrated directly. This channel can be indirectly activated by cholecystokinin and acetylcholine and its properties can to a large extent explain most previous microelectrode findings¹². The present study of single-channel currents in acinar

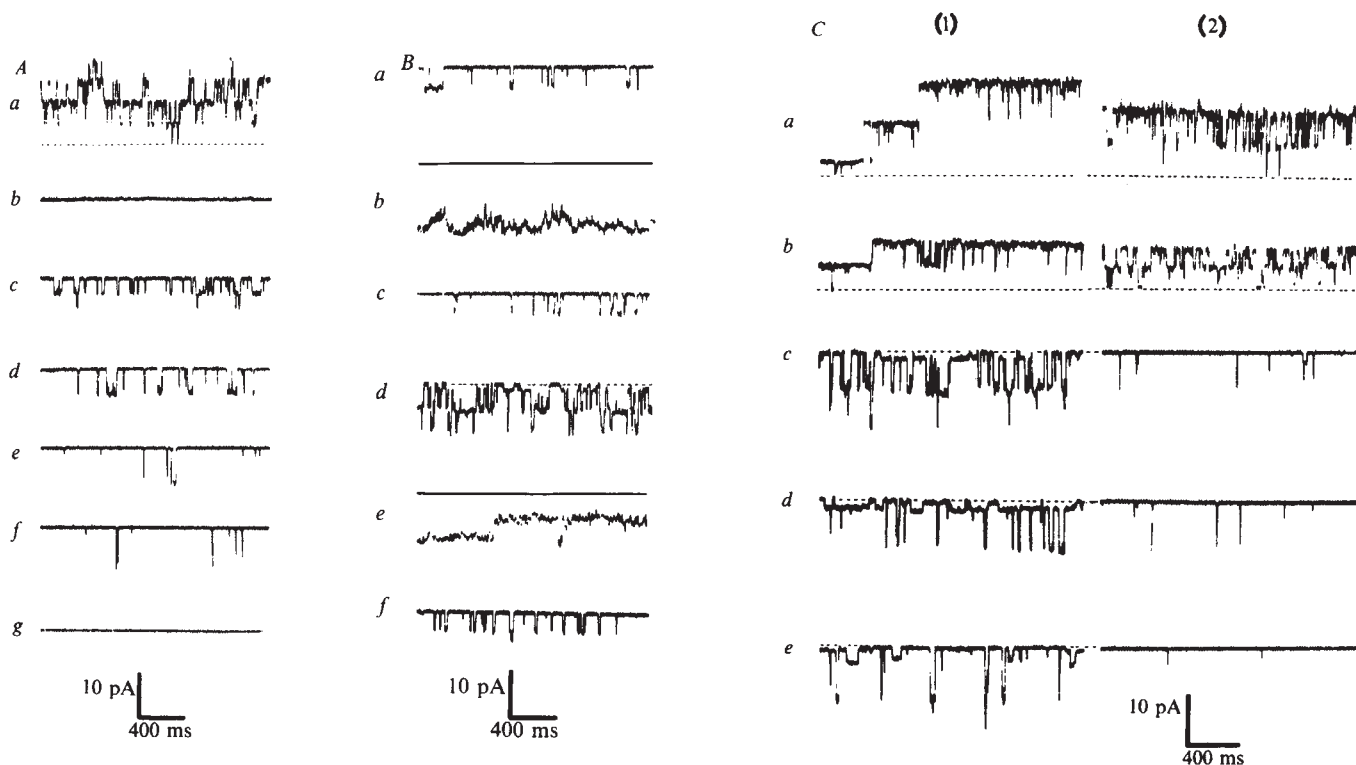


Fig. 4 Single-channel current traces from one mandibular acinar membrane patch. **A**, Recordings obtained *in situ* (cell-attached conformation). The pipette and bath solution was similar to the intracellular solution (see Fig. 2 legend) except that there was no EGTA and the Ca^{2+} concentration was 1.2 mM. *a-g* show consecutive traces obtained at different pipette potentials: *a*, -20 mV; *b*, 0 mV; *c*, +10 mV; *d*, +20 mV; *e*, +30 mV; *f*, +40 mV; and *g*, +50 mV. **B**, Recordings from the same patch as **A**, with *a*, still *in situ* (cell-attached conformation) and *b-f*, after excision (inside-out conformation). All traces were recorded using a pipette potential of +15 mV, but with different bath solutions (in contact with the intracellular aspect of the plasma membrane). *b*, Intracellular type solution but with 1.2 mM Ca^{2+} and no EGTA (same as pipette solution). *c*, Intracellular solution, that is, no Ca^{2+} added and 1 mM EGTA (internal Ca^{2+} concentration, $[\text{Ca}^{2+}]_i < 10^{-9}$ M). *d*, Intracellular solution with 1 mM Ca^{2+} and 2 mM EGTA ($[\text{Ca}^{2+}]_i \sim 10^{-7}$ M). *e*, Intracellular solution with 3 mM Ca^{2+} and 3.1 mM EGTA ($[\text{Ca}^{2+}]_i \sim 3 \times 10^{-6}$ M). *f*, Intracellular solution with no added Ca^{2+} and 1 mM EGTA ($[\text{Ca}^{2+}]_i < 10^{-9}$ M) (same as *c*). The horizontal lines (solid or broken) represent the current level when all channels were closed. **C**, Recordings from same excised inside-out patch as in **B**. Two different intracellular solutions were used in the bath: in (1) the solution contained 1 mM Ca^{2+} and 2 mM EGTA ($[\text{Ca}^{2+}]_i \sim 10^{-7}$ M), while in (2) there was no added Ca^{2+} and EGTA (1 mM) was present. *a-e* show traces obtained at different pipette potentials: *a*, -25 mV; *b*, -15 mV; *c*, +25 mV; *d*, +35 mV; and *e*, +45 mV.

cells from three different mammalian salivary glands now demonstrates the existence of a voltage-activated K^+ channel that had gone unnoticed in previous microelectrode studies^{1,14}. The properties of the channel seem similar to those described for the Ca^{2+} -dependent K^+ channels in the electrically excitable chromaffin cells⁸. The channel has a very high conductance (~ 250 pS in symmetrical K^+ solution), is very Ca^{2+} -sensitive (an increase of Ca^{2+} concentration from $< 10^{-9}$ M to 10^{-7} M has a dramatic effect), is activated by depolarization and is insensitive to changes in external Ca^{2+} concentration.

The membrane responses to nerve stimulation in the acinar cells are thought to be mediated by a rise in internal ionized Ca^{2+} concentration^{1,7,17}, the demonstration of a Ca^{2+} -activated K^+ channel provides direct evidence for this. The activation of the channel by membrane depolarization may also be physiologically relevant, as the immediate membrane effect of nerve stimulation in optimal recording conditions (microelectrode experiments) is depolarization¹⁸. The initial depolarization may be explained by the smaller (30–35 pS) Na^+ -permeable channel also observed in this study (Fig. 1E).

What is the physiological role of the K^+ release phenomenon? In some transporting epithelia (both secretory and absorbing) there is indirect evidence that coupled inward movement of Na^+ and Cl^- (Na^+ - Cl^- co-transport) occurs in the direction of the transepithelial transport¹⁹. This transport system has recently been shown to be a Na^+ - K^+ - Cl^- co-transporter^{20–22} and Musch *et al.*²² have shown that luminal K^+ stimulates NaCl absorption in the intestine of a marine teleost. Ion co-transport

systems do not seem to be directly controlled by extra or intracellular messengers, yet in many secretory tissues, the rate of salt and fluid transport can be increased rapidly by nerve stimulation. The voltage and Ca^{2+} -dependent K^+ channel, now demonstrated for the first time in transporting epithelia could, by allowing K^+ release in response to nerve stimulation, be indirectly responsible for activation of Na^+ - K^+ - Cl^- co-transport systems.

This work was supported by a MRC project grant.

Received 18 January; accepted 24 February 1983.

- Petersen, O. H. *The Electrophysiology of Gland Cells* (Academic, London, 1980).
- Burgen, A. S. V. *J. Physiol., Lond.* **132**, 20–39 (1956).
- Petersen, O. H. *J. Physiol., Lond.* **208**, 431–447 (1970).
- Nishiyama, A. & Petersen, O. H. *J. Physiol., Lond.* **242**, 173–188 (1974).
- Ginsborg, B. L., House, C. R. & Silinsky, E. *J. Physiol., Lond.* **236**, 723–731 (1974).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Ginsborg, B. L. & House, C. R. *Rev. Biophys. Bioengng* **9**, 55–80 (1980).
- Marty, A. *Nature* **291**, 497–500 (1981).
- Pallotta, B. S., Magleby, K. L. & Barrett, J. N. *Nature* **293**, 471–474 (1981).
- Barrett, J. N., Magleby, K. L. & Pallotta, B. S. *J. Physiol., Lond.* **331**, 211–230 (1982).
- Maruyama, Y. & Petersen, O. H. *Nature* **299**, 159–161 (1982).
- Maruyama, Y. & Petersen, O. H. *Nature* **300**, 61–63 (1982).
- Pedersen, G. L. & Petersen, O. H. *J. Physiol., Lond.* **234**, 217–227 (1973).
- Wakui, M. & Nishiyama, A. *Pflügers Arch. ges. Physiol.* **386**, 251–259 (1980).
- Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
- Yellen, G. *Nature* **296**, 357–359 (1982).
- Putney, J. W. *Pharmac. Rev.* **30**, 209–245 (1979).
- Gallagher, D. V. & Petersen, O. H. *J. Physiol., Lond.* **305**, 43–57 (1980).
- Frizzell, R. A., Field, M. & Schultz, S. G. *Am. J. Physiol.* **236**, F1–F8 (1979).
- Chipperfield, A. R. *Nature* **286**, 281–282 (1980).
- Greger, R. & Schlatter, E. *Pflügers Arch. ges. Physiol.* **392**, 92–94 (1981).
- Musch, M. W. *et al. Nature* **300**, 351–353 (1982).

Tyrosine hydroxylase activation in depolarized dopaminergic terminals—involvement of Ca^{2+} -dependent phosphorylation

S. El Mestikawy, J. Glowinski, & M. Hamon

Groupe NB, Inserm U.114, Collège de France, 11 Place Marcelin Berthelot, 75231 Paris Cedex 5, France

Tyrosine hydroxylase (tyrosine 3-monoxygenase, EC 1.14.16.2, TH) catalyses the rate limiting step of catecholamine biosynthesis. *In vitro*, TH from central dopaminergic¹⁻⁴ as well as from central^{5,6} and peripheral^{6,7} noradrenergic neurones can be activated by a cyclic AMP-dependent phosphorylation process and several authors⁷⁻⁹ have proposed that this process can be responsible for the *in vivo* activation of TH resulting from the electrical stimulation of these neurones. However, this is unlikely to be the case for TH in central dopaminergic neurones because depolarization produces an enzyme activation which is additive with that due to the cyclic AMP-dependent phosphorylation process¹⁰⁻¹². In the case of tryptophan hydroxylase in central serotonergic neurones, recent evidence indicates that a Ca^{2+} -dependent instead of a cyclic AMP-dependent phosphorylation process is responsible for the increased enzyme activity triggered by depolarization¹³. This finding led us to investigate whether a Ca^{2+} -dependent phosphorylation process also accounts for the activation of TH inside depolarized dopaminergic terminals. We found that soluble TH from the rat striatum could be activated by a Ca^{2+} -dependent process in optimal conditions for producing the phosphorylation of proteins. This activation corresponded exactly to that resulting from the incubation of striatal slices in K^+ -enriched medium and indeed TH activity from depolarized dopaminergic terminals could not be further stimulated by Ca^{2+} -dependent phosphorylating conditions. In contrast, *in situ* TH activation by cyclic AMP-dependent phosphorylation (triggered by dibutyryl cyclic AMP or forskolin) did not prevent subsequent stimulation by Ca^{2+} -dependent phosphorylation. These findings suggest that TH activation in depolarized dopaminergic terminals involves a Ca^{2+} -dependent phosphorylation process similar to that controlling tryptophan hydroxylase activity in serotonergic neurones.

The striata of adult male Charles River rats (250 ± 20 g) were sliced with a MacIlwain tissue chopper (thickness 0.3 mm) and incubated at 37°C for 10 min in 5 ml of a Krebs-Henseleit medium¹³ containing 0.1 mM EGTA instead of Ca^{2+} . Some incubations were carried out in the presence of 2 mM dibutyryl-cyclic AMP (Sigma) or 0.1 mM forskolin (Calbiochem), an activator of adenylate cyclase¹⁴. Depolarization was achieved by the addition of 0.1 ml of a mixture of 3 M KCl and 0.2 M CaCl_2 after 5 min of incubation. Experiments were stopped by filtering each sample through a nylon mesh (pore size 0.06 mm). Tissues on the nylon mesh were immediately frozen on dry ice and homogenized within the following hour in a buffer (final pH 7.0) composed of 20 mM Tris-acetate, 2 mM β -mercaptoethanol plus 10 mM NaF and 20 mM Na phosphate to inhibit phosphoprotein phosphatases¹⁵. For the studies on pH dependency of TH activity, tissues were homogenized in 0.01 M HEPES, pH 6.0, containing the same concentrations of β -mercaptoethanol, NaF and Na phosphate as those added to the Tris buffer. The enzyme activity was measured in the 40,000g supernatant of each homogenate with 10 μM tyrosine and 0.25 mM 6-methyl-5,6,7,8-tetrahydropterin (6-MPH₄, Sigma) usually at pH 7.0 as described elsewhere¹² except that 50 μM EGTA, 10 mM NaF and 20 mM Na phosphate were included in the standard assay mixture. Determinations of apparent K_m and V_{max} values were also made at pH 7.0 with six different

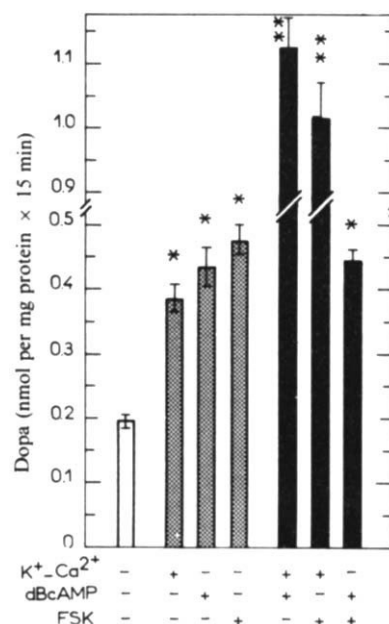


Fig. 1 Tyrosine hydroxylase activity in soluble extracts of striatal slices incubated in the presence of an excess of K^+ and/or dibutyryl cyclic AMP and/or forskolin. * $P < 0.05$ when compared with TH activity of control slices. ** $P < 0.05$ when compared with TH activity of slices incubated in the presence of 60 mM K^+ (plus 4 mM Ca^{2+}), 2 mM dibutyryl cyclic AMP or 0.1 mM forskolin alone. Incubating conditions are indicated under the figure: + = addition. dBcAMP, dibutyryl cyclic AMP; FSK, forskolin. Striatal slices were incubated at 37°C either for 10 min in Krebs-Henseleit medium containing 0.1 mM EGTA instead of Ca^{2+} or 5 min in this condition followed by 5 min in the same medium supplemented with 60 mM K^+ and 4 mM Ca^{2+} . Dibutyryl cyclic AMP (2 mM) or forskolin (0.1 mM) was added at the beginning of the incubation. Tissues (95 ± 5 mg) were then homogenized in 1.5 ml of 0.02 M Tris-acetate, pH 7.0, containing 2 mM β -mercaptoethanol, 10 mM NaF and 20 mM Na phosphate. TH activity was measured in an aliquot (0.2 ml) of the 40,000g supernatant of each homogenate. Each value is the mean $\pm$ s.e.m. of at least seven separate determinations.

concentrations of either tyrosine (5–100 μM) or 6-MPH₄ (0.1–2 mM). TH assays for assessing the pH dependence were performed in 0.05 M PIPES or HEPES buffer between pH 5.5 and 8.7.

As illustrated in Fig. 1, the addition of K^+ and Ca^{2+} (final concentrations 60 mM and 4.0 mM respectively) to the incubating medium doubled the activity of TH extracted from striatal slices. Both ions were necessary, as the addition of K^+ alone or of Na^+ (instead of K^+) plus Ca^{2+} did not affect the enzyme activity. Marked increases in TH activity were also observed when striatal slices were incubated in the presence of optimal concentrations of dibutyryl cyclic AMP (2 mM) or forskolin (0.1 mM) (Fig. 1). However, in contrast to that seen for K^+ -induced depolarization, neither agent required external Ca^{2+} to produce maximal activation of TH (Fig. 1). Kinetic analyses of TH indicated that the activation provoked by depolarization was associated with a marked increase in the V_{max} with no change in the apparent affinities of the enzyme for tyrosine and 6-MPH₄ (Table 1). The reverse was seen when TH was extracted from tissues incubated with dibutyryl cyclic AMP or forskolin, as the enzyme affinity for the pterin cofactor (but not for tyrosine) was the only parameter significantly increased (Table 1). Examination of the pH dependence of TH activity also revealed striking differences; only one optimum at pH 6.0 was observed with the enzyme from control or depolarized slices whereas two optima (pH 6.0 and 7.1) were found for that extracted from tissues incubated with dibutyryl cyclic AMP or forskolin (Table 1).

The similarities between changes in TH activity evoked by dibutyryl cyclic AMP and forskolin (Table 1) strongly suggested

that both compounds acted via the same mechanism, and the activation of TH by dibutyl cyclic AMP plus forskolin was no greater than that seen with only one of these agents (Fig. 1). In contrast, the stimulatory effect of K^+ -induced depolarization was increased in the presence of optimal concentrations of dibutyl cyclic AMP or forskolin (Fig. 1). Thus TH activation in slices exposed to 60 mM K^+ could not result from some depolarization-induced increase in cyclic AMP levels in tissues¹⁶.

As Ca^{2+} was an absolute requirement for the activation of TH in depolarized dopaminergic terminals, we looked for some effects of the cation on the enzyme itself. Ca^{2+} alone (20 μ M–500 μ M) in conditions provoking the activation of tryptophan hydroxylase by a Ca^{2+} -dependent neutral protease¹⁷ had no effect on TH activity in soluble striatal extracts. Similarly, Ca^{2+} (20–500 μ M) remained ineffective when added with $MgCl_2$ (5 mM) to the TH assay medium. In contrast, the addition of the Ca^{2+} -dependent phosphorylating mixture ATP (0.5 mM) + $MgCl_2$ (5 mM) + $CaCl_2$ (20–500 μ M) markedly increased TH activity in striatal extracts. This effect was blocked by trifluoperazine (IC_{50} 60 μ M), suggesting the involvement of calmodulin¹⁸. Indeed, calmodulin (2–5 μ g per assay) suppressed the inhibitory effect of trifluoperazine. The measurement of the incorporation of ^{32}P from [γ - ^{32}P]ATP into soluble proteins of the 40,000g supernatant of striatal homogenates revealed that

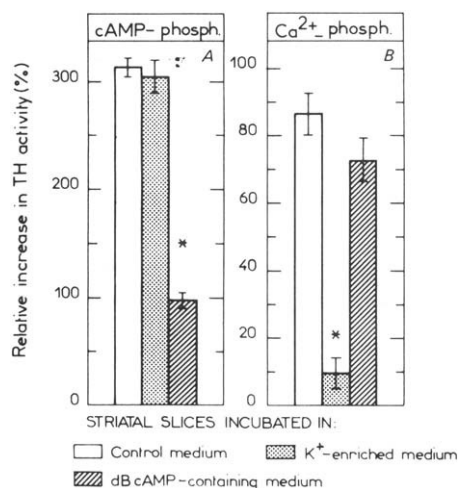


Fig. 2 Effects of cyclic AMP- or Ca^{2+} -dependent phosphorylation on the activity of TH from striatal slices incubated in control conditions, in the presence of an excess of K^+ or with dibutyl cyclic AMP. Striatal slices were incubated at 37°C either for 10 min in Ca^{2+} -free Krebs-Henseleit medium (open bars), for 10 min in the same medium supplemented with 2 mM dibutyl cyclic AMP (hatched bars) or for 5 min in the Ca^{2+} -free medium followed by 5 min after the addition of 60 mM K^+ and 4 mM Ca^{2+} (stippled bars). Tissues were homogenized in 0.02 M Tris-acetate, pH 7.0, supplemented with 2 mM β -mercaptoethanol, 10 mM NaF and 20 mM Na phosphate, and the 40,000g supernatants were filtered through Sephadex G25 columns (1.5 cm $\times$ 5.5 cm, PD 10 Pharmacia) equilibrated with the same buffer as that used for homogenization. TH activity was measured in the void volumes in standard assay conditions, in the presence of 0.5 mM ATP, 5 mM $MgCl_2$, 0.5 mM dibutyl cyclic AMP and 5 μ g of cyclic AMP-dependent protein kinase (from bovine heart, Sigma) (panel A) and in the presence of 0.5 mM ATP, 5 mM $MgCl_2$, 0.2 mM $CaCl_2$ and 10 μ g calmodulin (panel B). Each bar represents the % increase in TH activity due to the addition of the cyclic AMP (A) or the Ca^{2+} (B) dependent phosphorylating mixture to the assay. Basal TH activities were 0.221 ± 0.011 nmol 3H -dopa per mg protein $\times$ 15 min for slices incubated in control medium, 0.368 ± 0.016 nmol 3H -dopa per mg protein $\times$ 15 min for those in depolarizing conditions, and 0.415 ± 0.030 nmol 3H -dopa per mg protein $\times$ 15 min for slices incubated in the presence of dibutyl cyclic AMP (means $\pm$ s.e.m. of eight separate determinations). * $P < 0.05$ when compared with the relative increase observed with TH extracted from slices incubated in Ca^{2+} -free Krebs-Henseleit medium (open bars).

Ca^{2+} stimulated a phosphorylation process in conditions producing an activation of TH¹³. Furthermore, ^{32}P incorporation was affected by trifluoperazine and calmodulin in the same conditions as the activity of TH in the presence of ATP- Mg^{2+} - Ca^{2+} . These observations confirmed that TH activation by ATP, Mg^{2+} , Ca^{2+} and calmodulin was associated with a Ca^{2+} -dependent phosphorylation^{19,20}.

Kinetic analyses indicated that the activation of TH by ATP- Mg^{2+} - Ca^{2+} was associated with an increased V_{max} and no change in the enzyme affinities for tyrosine and 6-MPH₄ (Table 1). In contrast, the cyclic AMP-dependent phosphorylation produced a selective increase in the apparent affinity for 6-MPH₄ with no significant alteration of V_{max} (Table 1). Other differences between the two activating processes concerned the pH dependence of TH: Ca^{2+} -dependent phosphorylation did not change the optimal pH of the enzyme (6.0) whereas the cyclic AMP-dependent process induced a second optimum at pH 7.1 (Table 1).

When TH was first activated by the cyclic AMP-dependent phosphorylation process and filtered through a Sephadex G-25 column (to remove the phosphorylating mixture), it was still sensitive to the stimulatory effect of ATP- Mg^{2+} - Ca^{2+} and conversely, a previous activation by the Ca^{2+} -dependent mechanism did not alter the efficacy of the cyclic AMP-dependent phosphorylation process in increasing the enzyme activity (Table 1). These two phosphorylation mechanisms are thus distinct and probably involve separate sites on the enzyme molecule (or an enzyme and an effector).

As noted with TH in soluble extracts, consecutive activations by cyclic AMP-dependent phosphorylation process and ATP- Mg^{2+} - Ca^{2+} could be achieved using slices as the starting material. Thus, the enzyme extracted from slices exposed to dibutyl cyclic AMP or forskolin could be activated further by ATP- Mg^{2+} - Ca^{2+} (Fig. 2, Table 1). This activation was as pronounced as that observed with TH extracted from control tissues. In contrast, TH activation by the cyclic AMP-dependent phosphorylation process was markedly less with the enzyme extracted from striatal slices incubated with dibutyl cyclic AMP or forskolin than with that from control or depolarized tissues (Fig. 2, Table 1). This led to the conclusion that dibutyl cyclic AMP and forskolin triggered *in situ* TH activation by the same mechanism as that due to *in vitro* cyclic AMP-dependent phosphorylation of the enzyme.

Unlike TH from tissues incubated with dibutyl cyclic AMP (Fig. 2) or forskolin (Table 1), the enzyme from depolarized tissues was as responsive as that from control slices to subsequent activation by the cyclic AMP-dependent phosphorylation process; however, it lost its sensitivity towards the Ca^{2+} -dependent activating mechanism (Fig. 2, Table 1). Since the 40,000g supernatants of control and depolarized striatal slices promoted equally the incorporation of ^{32}P from [γ - ^{32}P]ATP into exogenous histones (from calf thymus, Sigma) in the presence of 5 mM $MgCl_2$, 0.2 mM $CaCl_2$ and calmodulin (5 μ g per assay, IBF) (not shown), the lack of effect of the ATP- Mg^{2+} - Ca^{2+} mixture on TH activity from depolarized tissues could not be ascribed to some defect in the phosphorylating process. Instead, the data in Fig. 2 and Table 1 strongly suggest that TH was no longer responsive simply because a Ca^{2+} -dependent phosphorylation process had already occurred *in situ*, in depolarized dopaminergic terminals.

In conclusion, the present findings confirmed that cyclic AMP-dependent phosphorylation induces an activation of TH associated with an increased enzyme affinity for the pterin cofactor and marked changes in the pH dependence of its activity^{1,3-12}. As the effect of dibutyl cyclic AMP (which, unlike cyclic AMP, is transported across cell membranes) could be reproduced by forskolin, an activator of adenylate cyclase in various cell types including neurones¹⁴, it is likely that adenylate cyclase is present in striatal dopaminergic terminals.

In agreement with other authors¹⁰⁻¹², we observed that the activation of TH in depolarized dopaminergic terminals was clearly distinct from that evoked by cyclic AMP-dependent

Table 1 Characteristics of TH activity in the soluble fraction of freshly dissected (I) or incubated (II) striatum

Preparations	I			II		
Assay (I) or incubation (II) conditions	Standard	ATP-Mg ²⁺ dBcAMP	ATP-Mg ²⁺ -Ca ²⁺	K ⁺ -Ca ²⁺	Forskolin	Control
K _m (Tyr) μ M	25.1 $\pm$ 2.5	30.1 $\pm$ 3.8	27.0 $\pm$ 3.1	28.8 $\pm$ 3.9	30.7 $\pm$ 2.5	30.8 $\pm$ 5.5
K _m (6-MPH ₄) mM	0.86 $\pm$ 0.15	0.27 $\pm$ 0.04*	1.11 $\pm$ 0.14	1.14 $\pm$ 0.13	0.30 $\pm$ 0.06**	0.99 $\pm$ 0.14
V _{max} (nmol ³ H-dopa per mg protein $\times$ 15 min)	3.68 $\pm$ 0.50	4.14 $\pm$ 0.39	6.17 $\pm$ 0.36*	5.74 $\pm$ 0.61**	3.57 $\pm$ 0.38	3.17 $\pm$ 0.37
Optimal pH	6.0	6.0-7.1	6.0	6.0	6.0-7.1	6.0
% Activation by ATP, Mg ²⁺ , dBcAMP (cAMP-dependent protein kinase)	288-353	21-33	272-403	258-330	48-61	249-334
% Activation by ATP, Mg ²⁺ , Ca ²⁺ (calmodulin)	79-98	59-82	0-11	0-12	58-75	68-90

The enzyme sources were the 40,000g supernatants of striata homogenized in 0.02 M Tris-acetate, pH 7.0, containing 2 mM β -mercaptoethanol, 10 mM NaF and 20 mM Na phosphate. Preparation I was obtained from freshly dissected striata and II from striatal slices incubated either in Ca²⁺-free Krebs-Henseleit medium ('Control'), in the same medium supplemented with 0.1 mM forskolin ('Forskolin') or in depolarizing conditions (60 mM K⁺ and 4 mM Ca²⁺). Activation by cyclic AMP-dependent phosphorylation was achieved by adding 0.5 mM ATP, 5 mM MgCl₂ and 0.5 mM dibutyl cyclic AMP to the assay mixture. Exogenous (from bovine heart) cyclic AMP-dependent protein kinase (5 μ g per assay) was added only for the activation of TH in preparations II. TH assays in Ca²⁺-dependent phosphorylating conditions were carried out in the presence of 0.5 mM ATP, 5 mM MgCl₂ and 0.2 mM CaCl₂. Calmodulin (10 μ g per assay) promoted the activation by ATP, Mg²⁺, Ca²⁺ only in preparations II. When the 40,000g supernatant (I) was first incubated in Ca²⁺ (or cyclic AMP)-dependent phosphorylating conditions¹³ and then assayed for TH activity in the presence of the cyclic AMP (or Ca²⁺)-dependent phosphorylating mixture, it was filtered through a Sephadex G-25 column (PD 10, Pharmacia) just before TH assays to remove the small molecules (ATP, Mg²⁺...) which were added for the first phosphorylation. Each K_m or V_{max} value is the mean $\pm$ s.e.m. of 5-7 separate determinations. Optimal pHs are the means of data obtained in three independent experiments. * $P < 0.05$ compared with the corresponding value of preparation in standard assay conditions. ** $P < 0.05$ compared with the respective 'Control' (preparation II) value. The activation of TH due to the addition of ATP, Mg²⁺, dBcAMP ($\pm$ cAMP-dependent protein kinase) or ATP, Mg²⁺, Ca²⁺ ($\pm$ calmodulin) to the assay mixture (left-hand column) is expressed as the per cent increase over basal activity measured with 10 μ M ³H-tyrosine and 250 μ M 6-MPH₄ (subsaturating condition). The lower and upper increases are indicated for each activating condition (determinations from 3-7 separate experiments). The dotted frame encloses the similar data obtained with TH in Ca²⁺-dependent phosphorylating conditions (preparation I) and with the enzyme extracted from striatal slices incubated in depolarizing conditions (K⁺-Ca²⁺, preparation II).

phosphorylation, corresponding exactly to that evoked by Ca²⁺-dependent phosphorylation (Table 1). Thus TH activation in depolarized dopaminergic terminals is primarily triggered by the influx of Ca²⁺ which promotes the activity of a Ca²⁺-dependent protein kinase phosphorylating the enzyme itself^{19,20} or one of its effectors. In this respect, the regulation of TH inside dopaminergic terminals closely resembles that previously reported for tryptophan hydroxylase inside serotonergic terminals¹³.

This research was supported by grants from INSERM, DGRST (80.70.184) and Rhône-Poulenc SA.

Auto-inhibition of brain histamine release mediated by a novel class (H₃) of histamine receptor

Jean-Michel Arrang, Monique Garbarg & Jean-Charles Schwartz

Unité de Neurobiologie U109, Centre Paul Broca de l'INSERM, 2ter, rue d'Alésia, 75014 Paris, France

Although histaminergic neurones have not yet been histochemically visualized, there is little doubt that histamine (HA) has a neurotransmitter role in the invertebrate and mammalian central nervous system^{1,2}. For example, a combination of biochemical, electrophysiological and lesion studies in rats have shown that histamine is synthesized in and released from a discrete set of neurones ascending through the lateral hypothalamic area and widely projecting in the telencephalon³⁻⁵. Histamine acts on target cells in mammalian brain via stimulation of two classes of receptor (H₁ and H₂) previously characterized in peripheral organs^{6,7} and probably uses Ca²⁺ and cyclic AMP, respectively, as second messengers⁸⁻¹⁰. It is well established that several neurotransmitters affect neuronal activity in the central nervous system through stimulation not only of postsynaptic receptors, but also of receptors located presynaptically which often display distinct pharmacological specificity and by which they may control their own release. Such 'autoreceptors' have been demonstrated (or postulated) in the case of noradrenaline, dopamine, serotonin, acetylcholine and γ -aminobutyric acid (GABA) neurones^{11,12} but have never been demonstrated for histamine. We show here that histamine inhibits its own release from depolarized slices of rat cerebral cortex, an action apparently mediated by a class of receptor (H₃) pharmacologically distinct from those previously characterized, that is, the H₁ and H₂ receptors.

Received 3 December 1982; accepted 15 February 1983.

- Lovenberg, W., Bruckwick, E. A. & Hanbauer, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2955-2958 (1975).
- Joh, T. H., Park, D. H. & Reis, D. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4744-4748 (1978).
- Vrana, K. E., Allhiser, C. L. & Roskoski, R. Jr *J. Neurochem.* **36**, 92-100 (1981).
- Pollock, R. J., Kapatos, G. & Kaufman, S. J. *Neurochem.* **37**, 855-860 (1981).
- Acheson, A. L., Kapatos, G. & Zigmond, M. J. *Life Sci.* **28**, 1407-1420 (1981).
- Weiner, N. *Aromatic Amino Acid Hydroxylases and Mental Disease* (ed. Youdim, M. B. H.) 141-190 (Wiley, Chichester, 1979).
- Weiner, N., Lee, F. L., Meligeni, J. & Tank, A. W. *Function and Regulation of Monoamine Enzymes: Basic and Clinical Aspects* (eds Usdin, E., Weiner, N. & Youdim, M. B. H.) 3-14 (Macmillan, London, 1981).
- Murrin, L. C., Morgenroth, V. H. III & Roth, R. H. *Molec. Pharmac.* **12**, 1070-1081 (1976).
- Roth, R. H. & Salzman, P. M. *Structure and Function of Monoamine Enzymes* (eds Usdin, E., Weiner, N. & Youdim, M. B. H.) 149-168 (Dekker, New York, 1977).
- Goldstein, M., Bronaugh, R. L., Ebstein, B. & Roberge, C. *Brain Res.* **109**, 563-574 (1976).
- Bustos, G. & Roth, R. H. *Biochem. Pharmac.* **28**, 3026-3028 (1979).
- Simon, J. R. & Roth, R. H. *Molec. Pharmac.* **16**, 224-233 (1979).
- Hamon, M., Bourgoin, S., Artaud, F. & Glowinski, J. *J. Neurochem.* **33**, 1031-1042 (1979).
- Seamon, K. B., Padgett, W. & Daly, J. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3363-3367 (1981).
- Yamauchi, T. & Fujisawa, H. *J. biol. Chem.* **254**, 6408-6413 (1979).
- Kakiuchi, S., Rall, T. W. & McIlwain, H. *J. Neurochem.* **16**, 485-491 (1969).
- Hamon, M., Bourgoin, S., Artaud, F. & Héry, F. *J. Neurochem.* **28**, 811-818 (1977).
- Levin, R. M. & Weiss, B. *Molec. Pharmac.* **12**, 581-589 (1976).
- Raese, J. D., Makk, G. & Barchas, J. D. *Function and Regulation of Monoamine Enzymes: Basic and Clinical Aspects* (eds Usdin, E., Weiner, N. & Youdim, M. B. H.) 105-114 (Macmillan, London, 1981).
- Yamauchi, T. & Fujisawa, H. *Biochem. biophys. Res. Commun.* **100**, 807-813 (1981).

We studied histamine release from slices of rat cerebral cortex previously labelled by preincubation in the presence of ^3H -histidine. The cerebral cortex was selected for study as histamine synthesis in this region occurs predominantly in terminals of extrinsic neurones as shown by its occurrence in synaptosomes¹³, its strong ipsilateral reduction following lesions at the level of the lateral hypothalamic area^{3,14} and its almost total interruption after complete cortical deafferentation¹⁵. Labelling of the endogenous histamine stores by the ^3H -labelled precursor¹⁶ instead of the ^3H -amine itself (as is performed in most *in vitro* release studies of catecholamines, for instance) was selected because an active re-uptake mechanism analogous to that operating for other amine or amino acid neurotransmitters has never been demonstrated for histamine in brain, which would allow its selective entry into histaminergic neurones. Thus, although brain slices incubated in the presence of ^3H -histamine are labelled and thereafter release the ^3H -amine when depolarized¹⁷⁻²⁰, several features make it doubtful that these processes involve a transport system operating selectively within histaminergic neurones: (1) the entry of the labelled amine into cerebral slices is extremely slow, equilibrium not being reached for several hours¹⁷; (2) the process is not saturable, does not occur against marked concentration gradients and its energy-dependent character has never been demonstrated (refs 17-19 and J.M.A., M.G. and J.C.S., unpublished

observations); (3) the extent to which slices from several brain regions are labelled by ^3H -histamine varies but does not parallel the distribution of histaminergic neurone markers like endogenous histamine content or L-histidine decarboxylase activity¹⁸.

Cortical slices were preincubated in the presence of $3.3 \times 10^{-7} \text{ M}$ ^3H -histidine and a series of washing steps were performed to eliminate excess ^3H -labelled precursor and to obtain a low constant basal efflux of ^3H -histamine. Addition of 30 mM K^+ consistently elicited in 2 min a fivefold increase in the level of ^3H -histamine in the medium which, in such conditions, represented $\sim 10\%$ of the total (Table 1). The magnitude of the depolarization-induced release, expressed as a percentage of the tissue content, was similar to that of endogenously synthesized ^3H -histamine¹⁶ or of endogenous histamine²¹ from hypothalamic slices. However, the K^+ -induced release was not accompanied by a stimulation of ^3H -histamine formation as that observed¹⁶ in slices to which the depolarizing stimulus was applied while high ^3H -histidine levels were still present in the medium; this suggests that, due to the extensive preliminary washings performed in the present experiments, the specific activity of the ^3H -precursor in the synthesis pools was reduced to negligible levels compared with those present during the 'synthesis period'. The effect of the 30 mM K^+ stimulus was totally abolished when slices were allowed to synthesize and release ^3H -histamine in a Ca^{2+} -free medium (Table 1a) or to release it in medium containing 10 mM Mg^{2+} , presumably acting as a Ca^{2+} antagonist (Table 1b). The calcium dependency of the depolarization-induced histamine release has also been reported for hypothalamic slices^{16,21}.

When 10^{-6} M nonradioactive histamine was present in the medium, the ^3H -histamine release elicited by 30 mM K^+ (over basal efflux) was inhibited by $\sim 50\%$, while the basal efflux itself was unaffected (Table 1c). The ^3H -histamine release induced by another depolarizing agent, veratridine, acting by promoting Na^+ permeability, was also significantly inhibited in the presence of 10^{-6} M exogenous histamine (Table 1e). It is unlikely that these effects of exogenous histamine resulted from an impaired synthesis of ^3H -histamine via end-product inhibition because (1) cerebral L-histidine decarboxylase activity in soluble extracts is not modified in the presence of histamine in concentrations as high as 10^{-2} M (refs 16, 22); (2) no significant ^3H -histamine formation was occurring in the slices at the time when the depolarizing stimuli were applied as shown by the unmodified tissue levels of ^3H -histamine in the presence of α -hydrazinohistidine, an L-histidine decarboxylase inhibitor^{16,23}; also the inhibitory effect of exogenous histamine persisted in the presence of the synthesis inhibitor (Table 1d). It is also unlikely that the action of exogenous histamine resulted from an isotopic dilution of the endogenously synthesized ^3H -histamine in the tissues, because an effect of similar magnitude was still observed when the nonradioactive amine was added together with the depolarizing stimulus (see Fig. 2 legend) or when ^3H -histamine was synthesized from ^3H -histidine of 250 times lower specific activity (not shown). In contrast, the addition of exogenous histamine does not modify the release process in slices directly labelled with ^3H -histamine²⁰, again indicating that the latter does not significantly enter the endogenous histamine pools. Finally, an inhibition of ^3H -histamine inactivation processes by exogenous histamine would have resulted in an opposite effect—that is, an increased recovery in the medium. In fact, the hypothesis that the inhibitory effect of exogenous histamine in slices is a receptor-mediated process is substantiated by observations that it displays characters of saturability, reversibility, high pharmacological specificity and that it is antagonizable in an apparently competitive manner.

The inhibitory action of histamine was concentration-dependent (Fig. 1), with a maximal inhibition of $61 \pm 3\%$ of K^+ -evoked ^3H -histamine release and a half-maximal effect of a $4.1 \pm 0.8 \times 10^{-8} \text{ M}$ concentration of histamine added 5 min before the K^+ stimulus (when histamine was added together with the K^+ stimulus, this EC_{50} value was increased threefold, while the

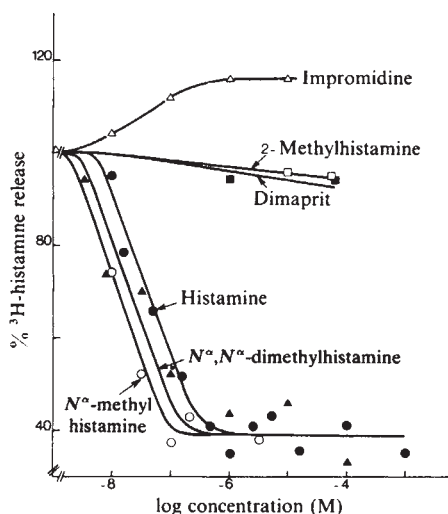


Fig. 1 Effect of various histamine agonists on the K^+ -evoked release of ^3H -histamine from slices of rat cerebral cortex. Rat cortical slices were prepared, preincubated with ^3H -histidine and washed by seven successive transfers as described in Table 1 legend. After the last washing period, $250\text{-}\mu\text{l}$ aliquots of the slice suspension ($2.5\text{--}3.5 \text{ mg protein}$) were distributed in plastic tubes (Eppendorf) containing $10\text{-}\mu\text{l}$ of the solutions of histamine or agonists (adjusted to $\text{pH } 7.0$) and kept at 37°C . Five minutes later, $250\text{-}\mu\text{l}$ of a modified Krebs-Ringer's solution were added to give a final concentration of either 2 mM or 30 mM K^+ and the incubations were stopped after 2 min by rapid centrifugation. The spontaneous efflux of ^3H -histamine into the medium in the presence of 2 mM K^+ represented $3.0 \pm 0.2\%$ of the total ^3H -amine present in tissue plus medium (that is, $3,648 \pm 242 \text{ d.p.m. per mg protein in tissue}$ and $111 \pm 9 \text{ d.p.m. per mg protein in medium}$). None of the added agents altered this spontaneous efflux (not shown). In the presence of 30 mM K^+ and the absence of added agents, the ^3H -histamine level in the medium represented $12.9 \pm 0.8\%$ of the total ($3,538 \pm 292 \text{ d.p.m. per mg protein in tissue}$ and $524 \pm 50 \text{ d.p.m. per mg protein in medium}$). The K^+ -evoked release was calculated in each experiment as the difference between mean ^3H -histamine levels in 30 mM K^+ and 2 mM K^+ media respectively (both expressed as percentages of total ^3H -histamine) and represented $9.9 \pm 0.8\%$ (mean $\pm \text{s.e.m. of } 25$ experiments). Results are expressed as percentages of this value, each point representing at least triplicate determinations in 1-10 different experiments. The maximal inhibition of 30 mM K^+ -evoked ^3H -histamine release by exogenous histamine, N -methyl-histamine and N,N' -dimethylhistamine was $61 \pm 3\%$.

maximal inhibition was of similar magnitude; see Fig. 2b). The effect of exogenous histamine was reversible as inhibition of release was no longer observed when slices were rapidly washed after exposure to the amine and before depolarization (not shown). Because the inhibitory action of exogenous histamine on release was not prevented in the presence of 2×10^{-7} M tetrodotoxin (to block propagation of action potentials in slices), was less pronounced when release was evoked by a stronger depolarization (66 mM K^+ or 20 μ M veratridine) and was also observed on cortical synaptosomes depolarized by K^+ (J.M.A., M.G. and J.C.S., in preparation), it is likely to result, as in the

case of other neurotransmitters, from a direct stimulation of autoreceptors^{11,12}.

The inhibitory action of histamine was mimicked by its two N^α and N^ω , N^α -methyl derivatives (Fig. 1) which showed the same maximal effect but with about three- and twofold higher relative potency, respectively, while they are slightly less potent than histamine at typical H_1 and H_2 receptors (Table 2). In contrast, N^ω -methylhistamine or N^γ -methylhistamine (a major histamine metabolite in brain) had no significant inhibitory activity and the same was true for compounds like 2-methylhistamine, 4-methylhistamine, 2-thiazolyethylamine

Table 1 Release of newly synthesized 3H -histamine and its inhibition by exogenous histamine in slices of rat cerebral cortex

Conditions		^3H -histamine in tissue	^3H -histamine in medium		
		(d.p.m. per mg protein)	(d.p.m. per 0.5 ml per mg protein)	(% of total)	
<i>a</i>	Krebs-Ringer's	2 mM K^+	4,940 $\pm$ 230	108 $\pm$ 7	2.1 $\pm$ 0.1
	Krebs-Ringer's		30 mM K^+ 4,490 $\pm$ 170	530 $\pm$ 59 (+ 391%)	10.6 $\pm$ 1.1
	Krebs-Ringer's without Ca^{2+}	2 mM K^+	10,270 $\pm$ 590	200 $\pm$ 12	1.9 $\pm$ 0.1
	Krebs-Ringer's without Ca^{2+}	30 mM K^+	8,780 $\pm$ 700	190 $\pm$ 15	2.1 $\pm$ 0.1*
<i>b</i>	Krebs-Ringer's	2 mM K^+	3,320 $\pm$ 190	106 $\pm$ 8	3.1 $\pm$ 0.4
	Krebs-Ringer's	30 mM K^+	3,190 $\pm$ 310	456 $\pm$ 58 (+ 330%)	12.5 $\pm$ 1.6
	Krebs-Ringer's with 10 mM Mg^{2+}	2 mM K^+	3,427 $\pm$ 550	110 $\pm$ 8	3.1 $\pm$ 0.8
	Krebs-Ringer's with 10 mM Mg^{2+}	30 mM K^+	3,405 $\pm$ 85	127 $\pm$ 8	3.6 $\pm$ 0.1†
<i>c</i>	Krebs-Ringer's	2 mM K^+	3,329 $\pm$ 136	123 $\pm$ 6	3.5 $\pm$ 0.1
	Krebs-Ringer's	2 mM K^+ + 10^{-6} M histamine	4,039 $\pm$ 367	158 $\pm$ 9	3.8 $\pm$ 0.5
	Krebs-Ringer's	30 mM K^+	2,921 $\pm$ 114	588 $\pm$ 41 (+ 378%)	16.4 $\pm$ 0.8
	Krebs-Ringer's	30 mM K^+ + 10^{-6} M histamine	3,238 $\pm$ 120	352 $\pm$ 24 (+ 186%)	9.8 $\pm$ 0.5*
<i>d</i>	Krebs-Ringer's with 0.1 mM α -hydrazinohistidine	2 mM K^+	3,180 $\pm$ 110	179 $\pm$ 13	5.3 $\pm$ 0.2
		30 mM K^+	2,340 $\pm$ 90	794 $\pm$ 57 (+ 344%)	25.3 $\pm$ 0.7
	Krebs-Ringer's with 0.1 mM α -hydrazinohistidine	30 mM K^+ + 10^{-6} M histamine	2,403 $\pm$ 299	499 $\pm$ 67 (+ 179%)	17.2 $\pm$ 1.1*
<i>e</i>	Krebs-Ringer's	Veratridine (0 μM)	2,770 $\pm$ 130	94 $\pm$ 5	3.3 $\pm$ 0.2
	Krebs-Ringer's	Veratridine (10 μM)	2,560 $\pm$ 170	247 $\pm$ 10 (+ 163%)	8.8 $\pm$ 0.7
	Krebs-Ringer's	Veratridine (10 μM) + 10^{-6} M histamine	2,960 $\pm$ 250	193 $\pm$ 17 (+ 105%)	6.1 $\pm$ 0.1*

Rats were killed by decapitation and slices (0.3 mm thick) were obtained from the cerebral cortex with a McIlwain tissue chopper. Pooled slices from two animals were washed and resuspended in 12 ml of modified Krebs-Ringer's bicarbonate medium (mM): 120 NaCl, 0.8 KCl, 2.6 $CaCl_2$, 0.67 $MgSO_4$, 1.2 KH_2PO_4 , 27.5 $NaHCO_3$, 10 glucose, pH 7.4, gassed with O_2/CO_2 (95:5). The slices ($\sim$ 12 mg protein per ml) were allowed to synthesize 3H -histamine during 30 min incubation at 37 °C (under a constant stream of O_2/CO_2 (95:5) in the presence of 3.3×10^{-7} M 3H -histidine (41.10^6 d.p.m. ml^{-1}). Before each experiment 3H -histidine (57.5 Ci $mmol^{-1}$; Amersham) was purified by ion-exchange chromatography¹³ so that its 3H -histamine content was less than 0.01%. After 30 min, slices were transferred to an open plastic cylinder with a nylon mesh fitted to the bottom as a small basket and washed to remove excess 3H -histidine and to obtain a constant spontaneous 3H -histamine efflux. For this purpose, the basket was successively transferred to seven beakers containing 25 ml of fresh Krebs-Ringer's solution at 37 °C in O_2/CO_2 (95:5) in which it was kept for periods of 4 min (first four washings) and 2 min (last three washings). After the last washing period, 250- μ l aliquots of the slice suspension (2.5–3.5 mg protein) were distributed into plastic tubes (Eppendorf) kept at 37 °C and containing 250 μ l of a modified Krebs-Ringer's solution to give a final concentration of 2 or 30 mM K^+ (a–d) (in the presence of 30 mM K^+ , NaCl concentration was decreased to compensate, so that isomolarity was maintained), or 10 μ M veratridine (e). When required, exogenous histamine (10^{-6} M) was present at the start of the incubation. After 2 min, incubations were stopped by rapid centrifugation, pellets were homogenized in 200 μ l 0.01 M HCl and 3H -histamine present in the pellet and supernatant was isolated by ion-exchange chromatography on Amberlite CG 50 columns¹³. Radioactivity was estimated by liquid scintillation spectrometry at 45% counting efficiency; at least 1,000 disintegrations were counted over a background of 16 c.p.m. In each experiment, 3H -histamine recoveries and 3H -histidine contaminations in the isolation procedure were evaluated. For this purpose, test 3H -labelled compounds were applied to columns and submitted to the same washing and elution procedures as the extracts of pellet or medium. 3H -histamine recoveries were $86 \pm 1\%$ and 3H -histidine contaminations were less than 0.01%; experimental data were corrected accordingly. In a typical experiment with 3 mg of tissue protein, the amount of total radioactivity in slices at the onset of stimulation was 4.5×10^5 d.p.m., indicating that about 95% of the initial 3H -histidine had been removed during the extensive washing procedure. The levels of total radioactivity in tissues or medium were not significantly modified after addition of the various agents. 3H -histamine represented $2.4 \pm 0.4\%$ of total radioactivity in the tissue before application of depolarizing stimuli. The amounts of 3H -histamine found in the medium in typical experiments were 370 ± 20 d.p.m. and $1,760 \pm 120$ d.p.m., in the presence of 2 mM and 30 mM K^+ , respectively. In a control experiment the 3H -labelled material released from the slices by a 50 mM K^+ stimulus and eluted from the columns was identified as authentic 3H -histamine by enzymatic transformation into chloroform-extractable 3H - N^α -methylhistamine under the action of semi-purified histamine- N -methyltransferase in the presence of 10 μ M S -adenosylmethionine as a methyl donor³⁷; recovery in the organic solvent phase was the same as that for a 1.2×10^{-9} Ci sample of authentic 3H -histamine run in parallel. a, A Ca^{2+} -free medium was used from the beginning to the end of the experiment (quadruplicate determinations). The increased 3H -histamine levels in this experiment probably reflect the decrease of release during the successive washing periods¹⁶. b, The final concentration of 10 mM $MgSO_4$ was present only during the 2 min incubation with 2 mM or 30 mM K^+ (triplicate determinations). c, Values represent the mean $\pm$ s.e.m. of triplicate determinations from 41 experiments except for 2 mM K^+ plus 10^{-6} M histamine (two experiments with quadruplicate determinations). d, The last washing period (10 min) and the final 2 min incubation period were conducted in the presence of 10^{-6} M α -hydrazinohistidine (quadruplicate determinations). The total 3H -histamine content (tissue plus medium) was $3,132 \pm 120$ d.p.m. per mg protein compared with $2,850 \pm 131$ d.p.m. per mg protein obtained in a parallel experiment in the absence of α -hydrazinohistidine. In contrast, in a control experiment (not shown) the addition of the L-histidine decarboxylase inhibitor 10 min before 3H -histidine resulted in a 93% decrease in 3H -histamine synthesis in the slices (780 versus 11,310 d.p.m.). e, One typical experiment with quadruplicate determinations. Similar data were obtained in two additional distinct experiments with 5 μ M veratridine (data not shown).

* $P < 0.01$; † $P < 0.05$ compared with the corresponding controls.

or dimaprit which, however, all display measurable agonist activity at either H_1 or H_2 receptors (Fig. 1 and Table 2). Interestingly, impromidine, a potent but partial histamine agonist at H_2 receptors²⁴, not only failed to mimic the inhibitory action of histamine (at a concentration between 5×10^{-10} M and 10^{-5} M) but increased by $\sim 20\%$ the K^+ -evoked release of 3H -histamine. This facilitatory action of impromidine, although of limited amplitude, was highly reproducible, concentration-dependent (with a half-maximal effect at $\sim 3 \times 10^{-8}$ M) and was not observed on the spontaneous efflux of 3H -histamine (2 mM K^+). This suggested that impromidine might have been acting as an antagonist towards endogenous histamine released by the K^+ stimulus, thus revealing a normal negative feedback modulation of histamine release, as is the case for various autoreceptor antagonists^{11,12}. Indeed, the inhibitory action of exogenous histamine (10^{-6} M) on release was progressively reversed in the presence of impromidine at increasing concentrations, with a 50% inhibitory concentration (IC_{50}) of $5.7 \pm 1.1 \times 10^{-7}$ M (Fig. 2a), about 20 times higher than the drug concentration responsible for the half-maximal facilitatory effect (Fig. 1). The surmountable character of the antagonism of exogenous histamine by impromidine was demonstrated by opposing the drug at fixed concentrations to the amine at increasing concentrations; there was a progressive rightward shift of the concentration-response curve to histamine, whose maximal inhibitory effect was, in all cases, the same—that is, about 60% inhibition of K^+ -evoked release (Fig. 2b). Schild plot analysis of these data was performed while neglecting the influence of endogenous histamine (inset of Fig. 2b) and gave a straight line with a slope of 1.32 ± 0.34 , compatible with impromidine acting as a competitive antagonist of exogenous histamine. The pA_2 value of impromidine was 7.5, in good agreement with the K_B value of 6.5×10^{-8} M (Table 2) which is derived from its IC_{50} value in the experiment of Fig. 2a, assuming a competitive antagonism²⁵ and again neglecting the influence of endogenous histamine. Also, a K_B value of 6.3×10^{-8} M for impromidine was found when the drug was opposed to a 10^{-4} M concentration of

histamine and its IC_{50} value (4.9×10^{-5} M) was introduced into the Cheng-Prusoff equation²⁵ (not shown). Finally, the impromidine concentration ($\sim 3.10^{-8}$ M) required for the half-maximal facilitatory effect (Fig. 1) was close to these apparent dissociation constants. Interpreting this small facilitation as resulting from blockade by impromidine of a feedback inhibition of 3H -histamine release mediated by the same class of receptors as those on which exogenous histamine acts, leads to the idea that the 'equivalent concentration' of endogenous histamine was rather lower than the EC_{50} value of exogenous histamine, that is, 4×10^{-8} M. This would account for an inhibitory effect being still observable when exogenous histamine was added (the 'brake' is not maximally triggered by the endogenous amine).

The consistency of apparent dissociation constants of impromidine in these various tests indicates that neglecting the probably complex influence of endogenous histamine (whose concentration in the vicinity of putative autoreceptors varies with time and is influenced by the antagonists) has probably not biased these evaluations to any great extent. Therefore, although the apparent dissociation constants of antagonists must be regarded, to a certain extent, as approximations (as, inherently, in any other autoreceptor system) they have allowed us to characterize the receptors mediating the feedback inhibition of histamine release.

While H_1 -antihistamines like mepyramine, cyclizine or chlorpheniramine were ineffective in concentrations at which they block H_1 receptors, several H_2 -antihistamines antagonized in a concentration-dependent manner the inhibition of release elicited by 10^{-6} M histamine (Fig. 2a and Table 2) and the antagonism was surmounted by 2×10^{-3} M histamine (not shown). Moreover, the competitive nature of the antagonism elicited by the H_2 -antihistamine burimamide (Fig. 2c) was checked by Schild plot analysis of the data, giving a straight line with a slope of 1.19 ± 0.73 (inset of Fig. 2c). Again the pA_2 value of 7.5 for this compound was in good agreement with the K_B value of 7×10^{-8} M derived from its IC_{50} value

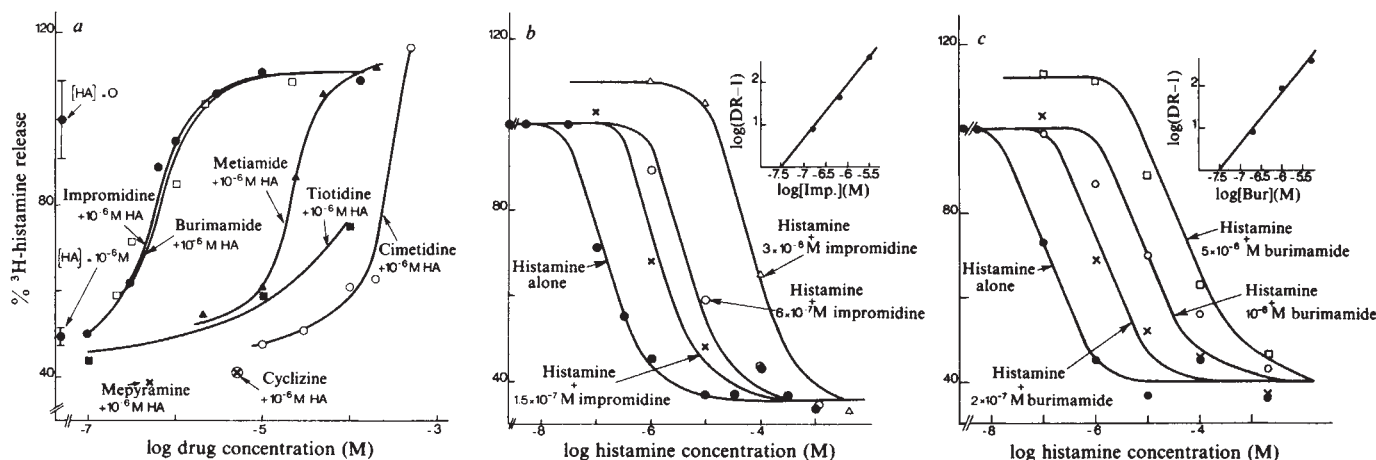


Fig. 2 Reversal by several drugs of the exogenous histamine-induced inhibition of K^+ -evoked 3H -histamine release. The procedure was that described in Fig. 1 legend except that exogenous histamine was added, when required, together with the depolarizing Krebs-Ringer's solution, 5 min after the various drugs (added in a 10- μ l volume of solution adjusted to pH 7). The spontaneous efflux of 3H -histamine (2 mM K^+) was $3.4 \pm 0.2\%$ ($3,213 \pm 154$ d.p.m. per mg protein in tissue and 115 ± 8 d.p.m. per mg protein in medium). In the presence of 30 mM K^+ and in the absence of added agents, the 3H -histamine level in medium represented $16.1 \pm 1.0\%$ of the total ($2,838 \pm 125$ d.p.m. per mg protein in tissue and 548 ± 48 d.p.m. per mg protein in medium). The results are expressed as percentages of the K^+ -evoked 3H -histamine release over spontaneous efflux, in the absence of exogenous histamine ($[HA] = 0$). a, The various drugs in increasing concentrations were opposed to exogenous histamine at a fixed concentration ($[HA] = 10^{-6}$ M) which, alone, inhibited the K^+ -evoked release by $50 \pm 2\%$ (mean $\pm$ s.e.m. of 15 experiments). None of the drugs, at the highest concentration tested, significantly modified spontaneous 3H -histamine efflux. b, c, Impromidine or burimamide at fixed concentrations were opposed to exogenous histamine at increasing concentrations. The insets represent the Schild plots of data, dose ratios (DR) being determined from the half-maximal inhibitory concentrations of histamine which were estimated graphically considering the total curve for each impromidine (Imp) or burimamide (Bur) concentration. b, The slope and 95% confidence limits for the regression of log(DR-1) on log[Imp] was 1.32 (0.98 – 1.66) and the coefficient of correlation was 0.998 . The pA_2 value for impromidine was 7.52 . c, The slope and 95% confidence limits for the regression of log(DR-1) on log[Bur] was 1.19 (0.47 – 1.92) and the coefficient of correlation was 0.990 . The pA_2 value for burimamide was 7.52 . Each point represents the mean of at least triplicate determinations in 1–14 different experiments.

Table 2 Comparison of potencies of histaminergic agents on the inhibition of cortical ^3H -histamine release, the guinea pig ileum contraction (H_1 receptors) and atrial rate (H_2 receptors)

Agent	Inhibition of ^3H -histamine release	Ileum contraction	Atrial rate
Relative agonist potency			
Histamine	100	100	100
N^+ -methylhistamine	<4	0.42	<0.1
N^+ -methylhistamine	<4	<0.01	<0.1
N^+ -Methylhistamine	270	72	74
N^+ , N^+ -dimethyl histamine	170	44	51
2-Methylhistamine	<0.08	16.5	4.4
2-Thiazolyethylamine	<0.008	26	2.2
4-Methylhistamine	<0.008	0.23	43
Dimaprit	<0.008	<0.0001	71
Impromidine	<0.03	<0.001	4,810
Antagonist activity (K_B , M)			
Mepyramine	$>5.8 \times 10^{-8}$	4.4×10^{-10}	
Cyclizine	$>5.8 \times 10^{-7}$	1.3×10^{-8}	
D-Chlorpheniramine	$>5.8 \times 10^{-8}$	5×10^{-10}	
L-Chlorpheniramine	$>5.8 \times 10^{-8}$	1.5×10^{-8}	
Metiamide	2.5×10^{-6}	$>10^{-3}$	9.2×10^{-7}
Cimetidine	3.3×10^{-5}	4.5×10^{-2}	7.9×10^{-7}
Burimamide	7×10^{-8}		7.8×10^{-6}
Ranitidine	$>1.2 \times 10^{-6}$		6.3×10^{-8}
Tiotidine	$>1.2 \times 10^{-5}$		1.5×10^{-8}
Impromidine	6.5×10^{-8}		4.8×10^{-7}
SKF 91486	8.8×10^{-8}		2.2×10^{-5}

Histamine and potential agonists were tested (in 2–11 different concentrations) as described in Fig. 1—that is, they were added 5 min before the K^+ stimulus. Their EC_{50} value was defined as the concentration resulting in a half-maximal inhibition of K^+ -evoked ^3H -histamine release, maximal inhibition being $61 \pm 3\%$ (see Fig. 1 legend). The relative potency was calculated as the ratio EC_{50} of histamine/ EC_{50} of agonist $\times 100$. The EC_{50} for histamine was $4.1 \pm 0.8 \times 10^{-8}$ M (values obtained by fitting the curve shown in Fig. 1 with an iterative computer least-squares method). Potential antagonists were added in at least four different concentrations 5 min before 10^{-6} M histamine and 30 mM K^+ and their IC_{50} s (concentration for which the inhibitory effect of exogenous histamine was reduced by 50%) determined in experiments similar to those of Fig. 2a using the iterative computer method³⁸. Apparent dissociation constants (K_B values) were calculated assuming a competitive antagonism, according to the equation²⁵: $K_B = \text{IC}_{50}/(1 + s/\text{EC}_{50})$, where s represents the concentration of exogenous histamine (10^{-6} M) and EC_{50} the amine concentration ($1.3 \pm 0.2 \times 10^{-7}$ M) eliciting a half-maximal inhibitory effect on K^+ -evoked release of ^3H -histamine (data from Fig. 2b). None of the agents altered the spontaneous efflux of ^3H -histamine at the highest concentration used to determine their antagonist activity. The values defining the activity of the various agents on the guinea pig ileum and atrium are taken from the review of Ganellin²⁶. Histaminergic agents were given by C. R. Ganellin. The following agents were inactive as antagonists: scopolamine, spiroperidol, naloxone, propranolol, mianserine, desipramine ($K_B > 10^{-7}$ M), phentolamine, domperidone ($K_B > 10^{-6}$ M) and imidazole ($K_B > 10^{-5}$ M).

against 10^{-6} M histamine (Table 2). Also, a facilitation of ^3H -histamine release similar to that elicited by impromidine was observed in the presence of high concentrations of burimamide (Fig. 2c) as well as of other antagonists (not shown). These effects of H_2 -antihistamines clearly cannot be attributed, however, to blockade of H_2 receptors as the potency of the various compounds markedly differed from that displayed at either peripheral or cerebral H_2 receptors^{26,27}. Thus, while the apparent dissociation constant of metiamide was in both cases in the micromolar range, large differences were observed for the other compounds: for example, burimamide⁷ which is 520 times less potent than tiotidine²⁸ at H_2 receptors was at least 170 times more potent than the latter at putative histamine autoreceptors. Similarly, SKF 91486 (3[4(5)-imidazolyl]propylguanidine), a weak H_2 receptor antagonist ($K_B = 2.2 \times 10^{-5}$ M)²⁹ was quite potent at these histamine receptors ($K_B = 8.8 \times 10^{-8}$ M). These differences do not arise from a restricted and variable diffusion of H_2 -antihistamines into the slice preparation because: (1) the apparent K_i values of the compounds were not modified when preincubation in their presence was extended from 5 to 15 min; (2) the K_B value of burimamide was lower (by two orders of magnitude) than on H_2 receptors; (3) in guinea pig hippocampal slices of similar thickness, the H_2 -antihistamines antagonize the stimulation of cyclic AMP formation mediated by H_2 receptors with potencies in close agreement

with those reported on other responses mediated by this homogeneous class of receptors (ref. 30 and J.M.A., M.G. and J.C.S., unpublished data).

Thus, from the relative potencies of histamine agonists and, more reliably³¹, from the apparent dissociation constants of histamine antagonists, as well as from the lack of effect of antagonists of other neurotransmitters (Table 2) it can be concluded that the auto-inhibition of histamine release in brain is mediated by an as yet unidentified class of histamine receptor; we propose calling these putative autoreceptors H_3 .

A striking feature of the histamine auto-inhibitory process is that it operates at concentrations of the amine at least 100 times lower than those required to stimulate putative post-synaptic receptors in the brain slice preparation, that is, H_1 receptors mediating glycogenolysis⁸ or H_2 receptors responsible for the stimulation of cyclic AMP accumulation^{9,32}. Although this is not a general feature, there are other examples in the literature of cerebral autoreceptors being more sensitive to the neurotransmitter than corresponding postsynaptic receptors: for example, noradrenaline inhibits its own release via α_2 adrenoreceptor stimulation at concentrations³³ approximately 100 times lower than those required to stimulate β -adrenoreceptors mediating the stimulation of cyclic AMP formation³⁴; a similar situation might exist for serotonin^{35,36}. It might be difficult at first sight to consider the means of functioning of a synapse in which the release of the neurotransmitter is already inhibited when its concentration in the cleft reaches levels well below the threshold for target cell stimulation. In fact, this is only an apparent paradox that can be resolved by taking into account several features of the system. First, even when maximally stimulated by the exogenous amine, autoreceptors inhibit the release only partially, that is, generally by $\sim 50\%$ as found for histamine, so that transmission can never be totally interrupted by this process. Second, the endogenous amine might not maximally trigger the auto-inhibitory system, even when its concentration in the cleft is largely above that required to saturate autoreceptors. Thus endogenous histamine maximally inhibited by 20% its own release (see the facilitatory action of impromidine in Fig. 1) although its concentration in the extracellular space of the slices was probably well above the micromolar range (estimation based on instantaneous release of 10% of the endogenous amine into the whole extracellular space of the slices). This incomplete auto-inhibition might result from spatial factors, the released amine not being able to reach before diffusion or inactivation all autoreceptors (like the extrasynaptic ones) that the exogenous amine stimulates. More likely, it might result from temporal factors: the concentration of endogenous histamine corresponding to maximal stimulation of autoreceptors may not be reached within the slices before the short delay beyond which it can no longer affect the release process. The physiological situation, indeed, clearly differs as the release is triggered by trains of impulses so that autoreceptor stimulation during a single impulse may affect the amount of neurotransmitter released by subsequent impulses.

In any event, the high pharmacological specificity of H_3 receptors which clearly differs from that of H_1 and H_2 receptors, suggests that development of compounds able to stimulate or block them selectively will provide useful tools for elucidating the precise functions of histaminergic neurones in the brain.

We thank Dr C. R. Ganellin and Smith-Kline & French Research for providing a fellowship (J.-M.A.).

Received 5 November 1982; accepted 28 February 1983.

- Weinreich, D. in *Biochemistry of Characterised Neurons* (ed. Osborne, C. N. N.) 153–175 (Pergamon, Oxford, 1977).
- Schwartz, J. C., Pollard, H. & Quach, T. T. *J. Neurochem.* **35**, 26–33 (1980).
- Garbarg, M., Barbin, G., Feger, J. & Schwartz, J. C. *Science* **186**, 833–835 (1974).
- Barbin, G., Garbarg, M., Schwartz, J. C. & Storm-Mathisen, J. *J. Neurochem.* **26**, 259–263 (1976).
- Haas, H. L. & Wolf, P. *Pflügers Arch. ges. Physiol.* **362**, 38 (1976).
- Ash, A. S. F. & Schild, H. O. *Br. J. Pharmac. Chemother.* **27**, 427–439 (1966).
- Black, J. W., Duncan, W. A. M., Durant, C. J., Ganellin, C. R. & Parsons, M. E. *Nature* **236**, 385–390 (1972).
- Quach, T. T., Duchemin, A. M., Rose, C. & Schwartz, J. C. *Molec. Pharmac.* **17**, 301–308 (1980).

9. Palacios, J. M., Garbarg, M., Barbin, G. & Schwartz, J. C. *Molec. Pharmac.* **14**, 971-982 (1978).
10. Hegstrand, L. R., Kanof, P. D. & Greengard, P. *Nature* **260**, 163-165 (1976).
11. Langer, S. Z. *Br. J. Pharmac.* **60**, 481-498 (1977).
12. Starke, K. A. *Rev. Pharmac. Tox.* **21**, 7-30 (1981).
13. Baudry, M., Martres, M. P. & Schwartz, J. C. *J. Neurochem.* **21**, 1301-1309 (1973).
14. Garbarg, M., Barbin, G., Bischoff, S., Pollard, H. & Schwartz, J. C. *Brain Res.* **106**, 333-348 (1976).
15. Barbin, G., Hirsch, J.-C., Garbarg, M. & Schwartz, J. C. *Brain Res.* **92**, 170-174 (1975).
16. Verdière, M., Rose, C. & Schwartz, J. C. *Eur. J. Pharmac.* **34**, 157-168 (1975).
17. Tuomisto, L., Tuomisto, J. & Walaszek, E. J. *Med. Biol.* **53**, 40-46 (1975).
18. Subramanian, N. & Mulder, A. H. *Eur. J. Pharmac.* **35**, 203-206 (1976).
19. Biggs, M. J. & Johnson, E. S. *Br. J. Pharmac.* **70**, 555-560 (1980).
20. Subramanian, N. *Life Sci.* **31**, 557-562 (1982).
21. Taylor, K. M. & Snyder, S. H. *J. Neurochem.* **21**, 1215-1223 (1973).
22. Schwartz, J. C., Lampart, C. & Rose, C. J. *J. Neurochem.* **19**, 801-810 (1972).
23. Schwartz, J. C., Lampart, C. & Rose, C. J. *J. Neurochem.* **17**, 1527-1534 (1970).
24. Durant, G. J. *et al. Nature* **276**, 403-405 (1978).
25. Cheng, Y. C. & Prusoff, W. H. *Biochem. Pharmac.* **22**, 3099-3108 (1973).
26. Ganellin, C. R. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 10-102 (Wright PSG, Bristol, 1982).
27. Schwartz, J. C. *et al. in Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 351-391 (Wright PSG, Bristol, 1982).
28. Yellin, T. O., Buck, S. H., Gilman, D. J., Jones, D. F. & Wardleworth, J. M. *Life Sci.* **25**, 2001-2009 (1979).
29. Parsons, M. E., Blakemore, R. C., Durant, G. J., Ganellin, C. R. & Rasmussen, A. C. *Ag. Actions* **5**, 464 (1975).
30. Trung Tuong, M. D., Garbarg, M. & Schwartz, J. C. *Nature* **287**, 548-551 (1980).
31. Black, J. W., Gerskowitz, V. P. & Leff, P. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 1-9 (Wright PSG, Bristol, 1982).
32. Dismukes, K., Rogers, M. & Daly, J. W. *J. Neurochem.* **26**, 785-790 (1976).
33. Wemer, J., Van Der Lugt, J.-C., De Langen, C. D. J. & Mulder, A. H. *J. Pharmac. exp. Ther.* **211**, 445-451 (1979).
34. Dolphin, A., Adrien, J., Hamon, M. & Bockaert, J. *Molec. Pharmac.* **15**, 1-15 (1979).
35. Mounsey, I., Brady, K. A., Carroll, J., Fisher, R. & Middlemiss, D. N. *Biochem. Pharmac.* **31**, 49-53 (1982).
36. Quach, T. T., Rose, C., Duchemin, A. M. & Schwartz, J. C. *Nature* **298**, 373-375 (1982).
37. Snyder, S. H., Baldessarini, R. J. & Axelrod, J. J. *Pharmac. exp. Ther.* **153**, 544-549 (1966).
38. Parker, R. B. & Waud, D. R. *J. Pharmac. exp. Ther.* **177**, 1-12 (1971).

Three loci related to the *src* oncogene and tyrosine-specific protein kinase activity in *Drosophila*

Michael A. Simon*, Thomas B. Kornberg* & J. Michael Bishop†

Department of Biochemistry and Biophysics* and Department of Microbiology and Immunology†, University of California, San Francisco, California 94143, USA

Rous sarcoma virus (RSV) is an acutely oncogenic avian retrovirus which induces sarcomas in animals and transforms fibroblasts in cell culture. Genetic analysis indicates that the viral *src* gene (*v-src*) mediates neoplastic transformation¹. The product of *v-src* is a 60,000 molecular weight (MW) phosphoprotein (pp60^{v-src}) possessing the enzymatic activity of a tyrosine-specific protein kinase²⁻⁶. The viral *src* gene is derived from a cellular gene (*c-src*) which also encodes a 60,000 MW phosphoprotein (pp60^{c-src}) with tyrosine-specific protein kinase activity⁴⁻¹⁰. Both birds and mammals are known to possess *c-src*^{7,8}. Shilo and Weinberg have reported that the genome of the fruit fly, *Drosophila melanogaster*, contains nucleotide sequences that are homologous to *v-src*¹¹. We report here the molecular cloning and chromosomal mapping of three loci from the *Drosophila* genome that contain such sequences. We also show that *Drosophila* contain both phosphotyrosine and a tyrosine-specific protein kinase activity immunoprecipitated by antisera directed against pp60^{v-src}. It should now be possible to identify the precise locus that encodes a *src*-specific protein kinase in *Drosophila*, and to explore the role of *c-src* in the growth and development of *D. melanogaster*.

A recombinant DNA library of the *Drosophila* genome cloned in bacteriophage λ (ref. 12) was screened by hybridization with a ³²P-labelled 800 base pair (bp) *PvuII* fragment of *v-src* (Fig. 1). Thirty positive clones were isolated and placed in one of three groups on the basis of analysis with restriction endonucleases. All positive clones were isolated several times. It is therefore unlikely that additional sequences that are homologous to *v-src* exist within the *Drosophila* genome. The approximate location of the nucleotide sequences within *v-src* that are homologous to the *Drosophila* clones was determined.

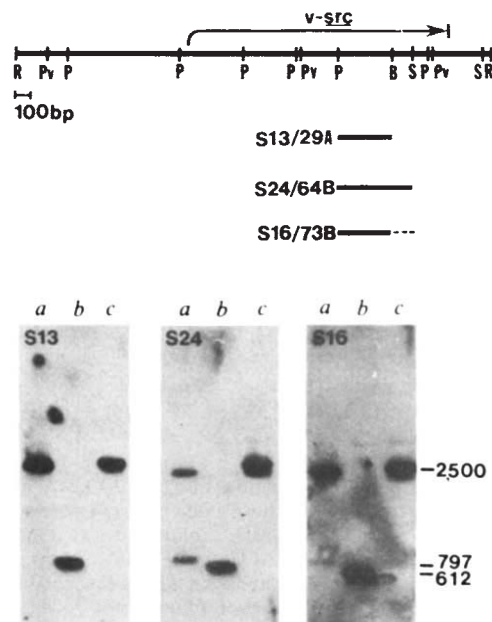


Fig. 1 Location within *v-src* of the sequences that hybridize to *Drosophila* clones S13, S24 and S16. The restriction map of the *EcoRI*-*B* fragment of the Schmidt-Ruppin A strain of RSV is shown (P = *PstI*, B = *BglI*, S = *SphI*, Pv = *PvuII*)²⁰. The *v-src* coding sequences are indicated by the arrow. A plasmid (1 μ g) containing the *EcoRI*-*B* fragment was digested with *EcoRI* and *BglI* (a), *EcoRI* and *PstI* (b) and *EcoRI* and *SphI* (c) and then electrophoresed and blotted onto nitrocellulose by the method of Southern¹³. The blot was probed with the nick-translated ³²P-labelled DNA of the *Drosophila* clone indicated on each panel and autoradiographed. Hybridization conditions were as described by Shilo and Weinberg with washing at 55 °C¹¹. The approximate sizes of the bands are indicated to the right of the autoradiograms. The extents of homology are summarized under the map of *v-src*. The dotted line indicates weak homology. Similar hybridization conditions were used for probing the *Drosophila* genomic library.

Restriction fragments from a plasmid containing the entire *v-src* gene were separated by electrophoresis through agarose gels, blotted from the gel onto nitrocellulose, and hybridized with ³²P-labelled DNA from different *Drosophila* clones¹³. Figure 1 shows the results using a single clone from each of the three groups. The patterns are distinct. Clone S13 hybridizes only with a 400 bp region defined by *PstI* and *BglI* sites. Clone S16 hybridizes strongly to the same 400 bp region and weakly to a 100 bp region defined by *BglI* and *SphI* sites. Clone S24 hybridizes strongly with both the 400 bp and 100 bp regions.

It is interesting that the region of *v-src* with which all three clones are homologous encodes a portion of the enzymatically active domain of pp60^{v-src} (ref. 14). This region of *v-src* also shows considerable homology to *v-fps*, the oncogene of Fujinami sarcoma virus, and *v-abl*, the oncogene of Abelson murine leukaemia virus (ref. 15 and D. Baltimore, personal communication). Both *v-fps* and *v-abl* are thought to encode tyrosine-specific protein kinases¹⁶⁻¹⁸. As indicated by Southern blot hybridization, *Drosophila* clones S16 and S13 show homology with *v-fps*, whereas clone S24 does not. Clone S16 also shows homology to *v-abl*, while S13 and S24 do not (data not shown).

Hybridization experiments with clones S13, S16 and S24 show little cross-hybridization among the three clones. This has allowed us to determine the chromosomal location of each clone without interference from related sequences in the other clones. *In situ* hybridization to *Drosophila* salivary chromosomes indicates that S13 maps to region 29A, S24 to region 64B and S16 to region 73B (Fig. 2).

If the genome of *Drosophila* contains genes that are functionally equivalent to *src*, *Drosophila* cells should possess tyrosine-specific protein kinase activity. We tested this

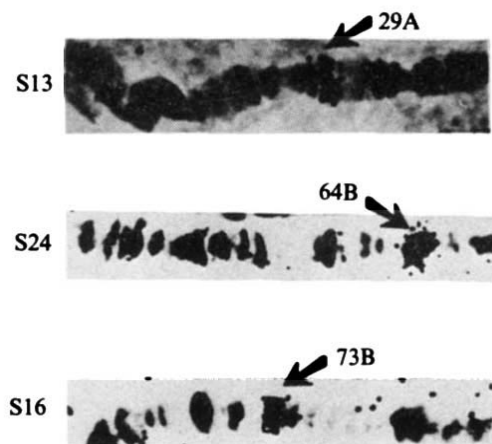


Fig. 2 The chromosomal locations of clones S13, S16 and S24. Salivary glands from third-instar Canton S larvae were squashed as described by Gall and Pardue²¹ onto slides that had been pretreated by the method of Brahic and Haase²². Squashes were heat-treated and prepared for hybridization as described by Bonner and Pardue²³. The hybridization solution was 300 mM NaCl, 30 mM sodium citrate, 30 mM sodium phosphate at pH 7.0, 40% formamide, 10% dextran sulphate, 300 $\mu\text{g ml}^{-1}$ sonicated salmon sperm DNA and 1×10^7 c.p.m. ml^{-1} of probe DNA that had been ^3H -labelled to a specific activity of 10^7 c.p.m. per μg DNA. Hybridization was for 16 h at 42°C under siliconized coverslips using 20 μl of hybridization solution per slide. Slides were washed, autoradiographed and stained as described by Gall and Pardue²¹.

prediction using antisera raised in newborn rabbits against RSV-induced tumours^{3,19}. All the tumour antisera immunoprecipitate pp60^{v-src}. The antisera do, however, have varying affinities for pp60^{c-src} (ref. 9). When an immune complex containing either pp60^{v-src} or pp60^{c-src} is incubated with ATP, a tyrosine of the immunoglobulin heavy chain is phosphorylated by the *src* protein^{2-6,9,10}. The results of such an immune complex kinase assay performed with extracts from Rat 2 cells, *Drosophila* K_c cells, and *Drosophila* embryos, larvae and adults are shown in Fig. 3a. Tumour serum 2 recovered detectable kinase activity from all of the extracts, whereas normal adult rabbit serum did not. Five of seven tumour sera tested detected kinase activity in K_c cells (data not shown). The two tumour sera which failed to detect kinase activity are specific to pp60^{v-src} and do not immunoprecipitate pp60^{c-src}. Phosphoamino acid analysis of the IgG chains phosphorylated by the K_c extract demonstrated that the phosphorylation was on a tyrosine residue (Fig. 3b).

The immune complex kinase assay is not a direct test of activity *in vivo*. We have, therefore, examined K_c cells for the presence of phosphotyrosine in order to demonstrate that tyrosine-specific protein kinases actually function in *Drosophila* cells. Figure 3c shows a two-dimensional phosphoamino acid analysis of ^{32}P -labelled K_c cells. Phosphotyrosine represented 0.1% of total phosphoamino acids. A value of 0.01% has been determined for Rat 2 cells using the same methods and is typical of untransformed cells (ref. 4 and our unpublished results). *Drosophila* cells must therefore contain tyrosine-specific protein kinases.

We have not yet been able to attribute the tyrosine-specific protein kinase activity in our immunoprecipitates directly to either a particular polypeptide or any of the three loci that contain homology to *v-src*. However, the tumour antisera used in this study do not react with tyrosine-specific protein kinases other than pp60^{v-src} and pp60^{c-src}. It is therefore likely that the tyrosine-specific protein kinase in *Drosophila* cells that is immunoprecipitated by these antisera is encoded by at least one of the three loci that we have cloned. We are now examining adult flies that are aneuploid for these chromosomal regions for gene dosage dependence of immune complex kinase activity.

Much of what is known about the role of pp60^{v-src} in neoplastic transformation has been gained either by studying cells before

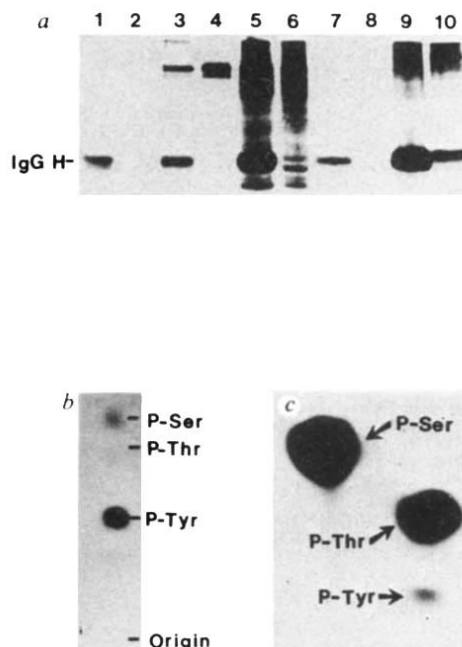


Fig. 3 *Drosophila* possess tyrosine-specific protein kinase activity. Immune complex kinase assays were performed as described previously³. Extracts were made from Rat 2 and K_c cells as described elsewhere³. Extracts from *Drosophila* embryos, larvae and adults were prepared in the same manner except that the samples were disrupted in a Dounce homogenizer. Approximately equal amounts of crude extract protein were assayed in the reactions represented by individual lanes in a. The position of the immunoglobulin heavy chain (IgG-H) was visualized by staining with Coomassie blue. The lanes are: 1, Rat 2 cells with tumour serum 2; 2, Rat 2 cells with normal serum; 3, K_c cells with tumour serum 2; 4, K_c cells with normal serum; 5, embryos with tumour serum 2; 6, embryos with normal serum; 7, larvae with tumour serum 2; 8, larvae with normal serum; 9, adults with tumour serum 2; 10, adults with normal serum. The IgG-H band from lane 3 was analysed for phosphoamino acid content as described by Hunter and Sefton⁴ except that following hydrolysis phosphoamino acids were purified by ion exchange²⁴ and electrophoresed on cellulose thin-layer plates at pH 3.5 for 90 min at 600 V. The plates were dried and autoradiographed for 1 day before the detection of markers by ninhydrin staining. The result is shown in b. c Shows a two-dimensional phosphoamino acid analysis of K_c cells. Approximately 10^7 cells were labelled for 3 h with 5 mCi of radioactive orthophosphate in 300 μl of D-20²⁵ media lacking phosphate. The cells were lysed in 0.3% SDS, 1% 2-mercaptoethanol, 50 mM Tris (pH 7.4) 5 mM MgCl_2 , 100 $\mu\text{g ml}^{-1}$ DNase I and 50 $\mu\text{g ml}^{-1}$ RNase A. After 10 min at 0°C , the proteins were precipitated by the addition of trichloroacetic acid to 15%. The proteins were hydrolysed and analysed as described by Cooper and Hunter²⁴. Markers were detected by ninhydrin staining. The radioactive spots were scraped into scintillation vials and counted in a toluene-based fluor.

and after RSV-induced transformation or by studying conditional mutants of *v-src*. The role of pp60^{c-src} in normal cellular physiology has remained an enigma largely due to the absence of similar genetic approaches. The presence of *c-src* and tyrosine-specific protein kinases in *Drosophila* may facilitate genetic analysis of the role of *c-src* in normal cells.

We thank M. Privalsky, G. Ramsey, M. Snyder and G. Mardon for helpful discussions, D. Baltimore for *v-abl* DNA, and J. Marinos for typing the manuscript. This work was supported by grants from the American Cancer Society and NIH.

Received 30 December 1982; accepted 2 March 1983.

- Vogt, P. K. in *Comprehensive Virology* Vol. 9 (eds Fraenkel-Conrat, H. & Wagner, R. R.) 341-455 (Plenum, New York, 1977).
- Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
- Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15**, 561-572 (1978).
- Hunter, T. & Sefton, B. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Collett, M. S., Purchio, A. F. & Erikson, R. L. *Nature* **285**, 167-169 (1980).

6. Levinson, A. D., Oppermann, H., Varmus, H. E. & Bishop, J. M. *J. biol. Chem.* **255**, 11973-11980 (1980).
7. Stehlin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170-173 (1976).
8. Spector, D., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4102-4106 (1978).
9. Oppermann, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
10. Collett, M. S., Erikson, E., Purchio, A. F., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3159-3163 (1979).
11. Shilo, B.-Z. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6789-6792 (1981).
12. Maniatis, T. *et al. Cell* **15**, 687-701 (1978).
13. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
14. Levinson, A. D., Courtneidge, S. A. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1624-1628 (1981).
15. Shibuya, M. & Hanafusa, H. *Cell* **30**, 787-795 (1982).
16. Pawson, T. *et al. Cell* **22**, 767-775 (1980).
17. Hanafusa, T., Mathey-Prevot, B., Feldman, R. A. & Hanafusa, H. *J. Virol.* **38**, 347-355 (1981).
18. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826-831 (1980).
19. Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-347 (1977).
20. DeLorbe, W. J., Luciw, P. A., Goodman, H. M., Varmus, H. E. & Bishop, J. M. *J. Virol.* **36**, 50-61 (1980).
21. Gall, J. G. & Pardue, M. L. *Meth. Enzym.* **210**, 470 (1971).
22. Brahic, M. & Haase, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6125-6129 (1978).
23. Bonner, J. J. & Pardue, M. L. *Chromosoma* **58**, 87-99 (1976).
24. Cooper, J. A. & Hunter, T. *Molec. cell. Biol.* **1**, 165-178 (1981).
25. Echalié, G. & Ohanessian, A. *In Vitro* **6**, 162-172 (1970).

Dispersion of the *ras* family of transforming genes to four different chromosomes in man

Stephen J. O'Brien*, William G. Nash†, Jerome L. Goodwin‡, Douglas R. Lowy§ & Esther H. Chang‡

* Laboratory of Viral Carcinogenesis, National Cancer Institute, NIH, and † Biological Carcinogenesis Program, NCI-FCRF, Frederick, Maryland 21701, USA

‡ Department of Pathology, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20014, USA

§ Dermatology Branch, National Cancer Institute, NIH, Bethesda, Maryland 20205, USA

Cellular transforming genes (*c-onc*) are evolutionarily conserved vertebrate DNA segments which have been identified by two different approaches. One group of these cellular genes has been defined by their close homology to the transforming genes of the acute transforming retroviruses (*v-onc*)¹⁻³. The second group, which represent activated forms of normal cellular genes⁴⁻⁹, has been detected by the ability of certain genes from animal and human tumours to induce focal transformation of tissue culture cells. Investigation of the possibility that the same cellular gene might have given rise to both a retroviral and a tumour transforming gene revealed that two of the *c-onc* genes identified by transfecting genomic DNA from human tumours to murine 3T3 fibroblasts were related to the transforming genes of two closely related acute transforming retroviruses, Harvey murine sarcoma virus (HaMuSV) and Kirsten murine sarcoma virus (KiMuSV)¹⁰⁻¹². The transforming genes of HaMuSV and KiMuSV are derived from two members of a cellular *onc* gene family called *ras*, which is a rather divergent group of normal vertebrate genes originally found by analysis of the cellular homologues of the *v-onc* genes of HaMuSV and KiMuSV¹³. Four distinct human cellular homologues of *v-Ha-ras* and *v-Ki-ras* (designated *c-Ha-ras* and *c-Ki-ras*, respectively) have been characterized¹⁴; two (*c-Ha-ras-1* and *c-Ha-ras-2*) are more closely related to *v-Ha-ras*, while the others (*c-Ki-ras-1* and *c-Ki-ras-2*) are more closely related to *v-Ki-ras*. On ligation with a retroviral long terminal repeat, the *c-Ha-ras-1* gene of both rat and human have been shown to induce *in vitro* transformation of mouse NIH 3T3 cells by DNA transfection^{15,16}. This gene and *c-Ki-ras-2* have also been isolated as activated transforming genes in human tumours¹⁰⁻¹². An understanding of the genetic relationship of the *c-ras* genes and additional genetic loci possibly involved in neoplastic transformation would be greatly facilitated by placement of the *ras* genes on the human chromosome map. Using DNA analysis of rodent × human somatic cell hybrids, we have now assigned each of the human genes to a different chromosome.

Somatic cell hybrids were constructed between fresh human lymphocytes (LLL) and rodent cells (mouse RAG and Chinese hamster E36) which were mutant in their hypoxanthine phosphoribosyl transferase gene, permitting hybrid selection on hypoxanthine-aminopterin-thymidine medium. Nine independent fusions were performed using PEG 1000, and 249 hybrids were derived (78 RAG × LLL and 161 E36 × LLL). These hybrids retained the entire rodent genome but segregated human chromosomes in different combinations. Following isozyme analysis of each of these primary hybrids¹⁷, two mapping panels (16 hybrids with RAG and 30 hybrids with E36) were chosen, selecting hybrids with low numbers of human chromosomes, but with strong isozyme signals for the human enzymes. For a given hybrid, high molecular weight DNA, isozyme extracts and karyotypic spreads were prepared at the same cell passage. G-11 chromosome staining was performed on each hybrid in the panel to discover hybrids having numerous chromosome rearrangements. This procedure stains human chromosomes light blue and rodent chromosomes magenta¹⁸. Hybrids with numerous human chromosome rearrangements or interspecific chromosome translocations were discarded. Each hybrid in the panel was analysed for up to 36 isozyme markers previously mapped to human chromosomes¹⁷. In addition, each hybrid was G-banded and the human chromosome complement determined¹⁹.

The *c-Ha-ras-1* gene was visualized as a 2.9-kilobase (kb)-fragment in human DNA following digestion with *SacI* (Fig. 1a), electrophoresis in 0.6% agarose gels, transfer to nitrocellulose and hybridization to a nick-translated probe derived from a molecularly cloned 2.9-kb human *SacI* fragment which includes the human *c-Ha-ras-1* locus¹⁴. A *SacI* digestion of E36 DNA did not produce any hybridization in this region of the filter. A lower stringency hybridization to a *SacI* digest using the same probe revealed an additional 12-kb fragment in human DNA (*c-Ha-ras-2* see below) and the Chinese hamster *c-Ha-ras* homologue (2.3 and 0.5 kb) in E36 DNA. Thus, by *SacI* digestion of DNAs from the hybrid panel, it was possible to determine which hybrids contained the human *c-Ha-ras-1* locus. The presence of *c-Ha-ras-1* showed perfect concordance with the presence of human chromosome 11, with *LDHA* (lactate dehydrogenase) and with *ACP2* (acid phosphatase-2), isozyme markers previously mapped to chromosome 11 (Table 1). Each of the remaining 22 human chromosomes showed high discordance with *c-Ha-ras-1* ($\geq 38\%$). These data, which indicate that *c-Ha-ras-1* is localized on human chromosome 11, independently confirm this assignment already made by others^{20,41}.

A 3.6-kb *BamHI* fragment characteristic of the *c-Ha-ras-2* was detected in Southern transfer of human DNA by hybridization to a nick-translated probe derived from a cloned 0.7-kb *BamI* fragment of human *c-Ha-ras-2* (Fig. 1b). This fragment contained nearly the entire *c-Ha-ras-2*-specific segment and approximately 100 additional bases of human flanking DNA which are not homologous to *v-Ha-ras* (see Fig. 1b). A low stringency hybridization of a *BamHI* digest of human DNA detects *c-Ha-ras-1* (6.6 kb, see Fig. 2a, map) but does not detect the *Ki-ras* loci¹⁴. A *BamHI* digest of E36 DNA revealed three bands (3.9, 4.5 and 8.6 kb) which did not co-migrate with the human signal, permitting detection of human *c-Ha-ras-2* in the hybrids. The *c-Ha-ras-2* gene was 100% concordant with the X chromosome and discordant ($\geq 40\%$) with the other 22 chromosomes (Table 1). These results permit the assignment of *c-Ha-ras-2* to the human X chromosome. One of the hybrids, 81P1, was positive for both *G6PD* (glucose 6-phosphate dehydrogenase) and *HPRT*, but negative for *c-Ha-ras-2* and for the X chromosome by banding. As these two markers are both located near the terminus of the long arm of the human X chromosome (*G6PD*-q28; *HPRT*-q26-q28, it seems that a centromere-proximal (to q26) region of the X contains *c-Ha-ras-2*. This region was lost in hybrid 81P1 and the X terminus (q26-q28) has been translocated to another chromosome, which we could not detect in our chromosome analyses.

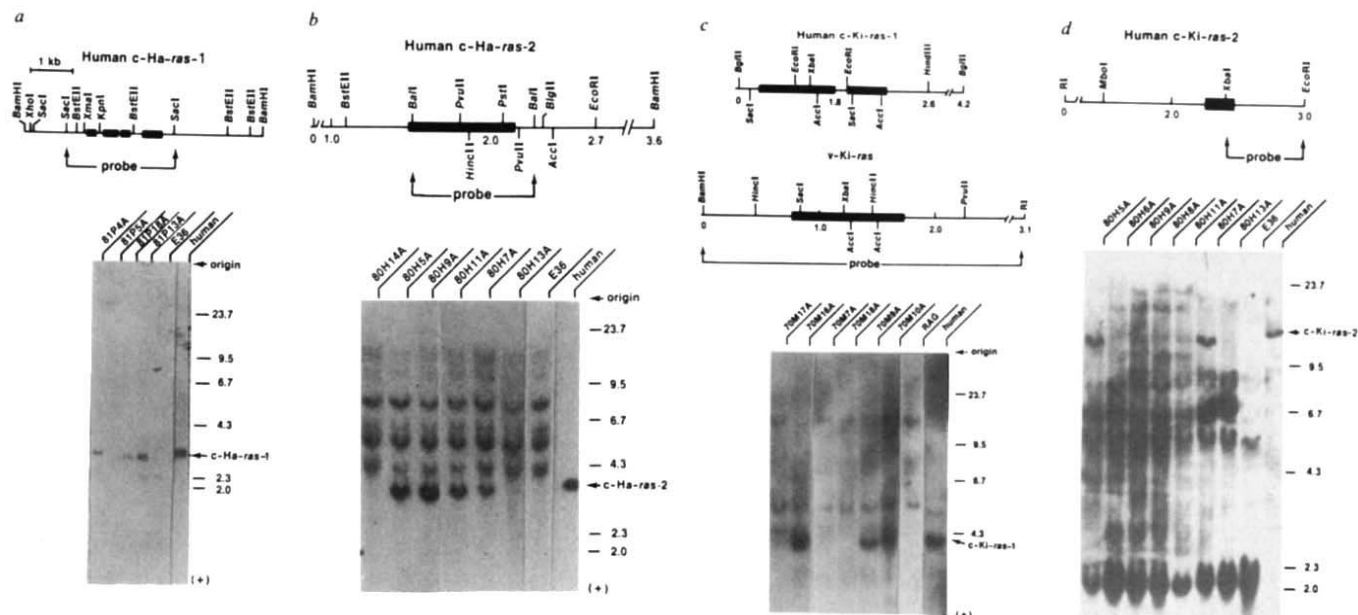


Fig. 1 Hybridization of labelled molecularly cloned *ras* probes to restriction enzyme digests of human, RAG mouse, E36 Chinese hamster and selected hybrids (see text and Table 1). Digestion was with *SacI* (a), *BamHI* (b), *BglII* (c) and *HindIII* (d). Restriction endonuclease-digested genomic DNA was electrophoresed in agarose gels, transferred to nitrocellulose filters and hybridized to various ^{32}P -labelled *ras* probes in a solution containing 0.60 M NaCl, 0.06 M sodium citrate, 0.2% Ficoll, 0.2% bovine serum albumin, 0.2% polyvinylpyrrolidone, 0.1% SDS and low molecular weight salmon testes DNA at $50\text{ }\mu\text{g ml}^{-1}$ at 65°C for 36–48 h¹⁴. Fragment sizes are given in kb. The restriction map of the human DNA segment which includes the *ras* homologues is presented above the gels. The hybridization probes are indicated as fragments which were isolated electrophoretically before nick-translation.

The *Ki-ras* genes are distantly related to the *Ha-ras* genes and in general the *Ha-ras* genes were not observed in Southern blots of human DNA by the specific *Ki-ras* probes used by us. The c-*Ki-ras-1* locus was detected as a 4.0-kb band in a *BglII* digest of human DNA using a 3.1-kb *BamHI*–*EcoRI* segment of v-*Ki-ras* (Fig. 1c). Chinese hamster DNA gave multiple bands using *BglII* and several other enzymes²¹, which made scoring of the c-*Ki-ras-1* locus in E36 $\times$ LLL hybrids ambiguous. *BglI* digestion of mouse cells exhibited only three fragments (16, 5.5 and 5.0 kb) none of which interfered with resolution of the human c-*Ki-ras-1* gene (Fig. 1c). The results of the RAG $\times$ LLL crosses are presented in Table 1 and illustrated in Fig. 1c. The presence of c-*Ki-ras-1* was concordant with human chromosome 6 and discordant ($\geq 23\%$) with each of the other chromosomes. We conclude that c-*Ki-ras-1* is localized on human chromosome 6.

The molecular probe which unequivocally identified c-*Ki-ras-2* was a cloned 0.5-kb *XbaI*–*EcoRI* fragment of human DNA which consists of approximately 100 bases or less of coding sequence homologous to v-*Ki-ras* and 400 bases of intervening sequence which is unique to this segment^{14,22}. This probe reveals a unique 14-kb fragment on hybridization to a *HindIII* digestion of human DNA (Fig. 1d). In similar hybridization conditions the hamster DNA showed at least four fragments of different sizes (8.0, 6.6, 5.6 and 2.3 kb) from the human segments. Of 40 hybrids scored, all showed concordance between c-*Ki-ras-2* and human chromosome 12 (Table 1). As all the other human chromosomes had large discordance with c-*Ki-ras-2*, we can conclude that c-*Ki-ras-2* is located on human chromosome 12.

We have previously found that two prototype loci, *Ha-ras* and *Ki-ras*, are present in at least two forms each in the normal human genome¹⁴. A fifth possible member of this family was recently isolated from a human neuroblastoma; however, its chromosomal position is unknown²³. The finding of four members of the *ras* gene family on four distinct chromosomes is perhaps surprising, but is not without precedent. The genes which encode the α - and β -subunits of haemoglobin have been mapped to different chromosomes in mammals^{24–28}. The transcriptionally active α - and β -globin genes are clustered, while the pseudogenes are dispersed to other chromosomes^{28,29}. The non-*ras* cellular transforming genes are also dispersed throughout the human genome (reviewed in ref. 30).

The c-*Ha-ras-1* genes of rats and humans have a similar genomic structure. The exons of these genes and a small portion of the introns form heteroduplexes, suggesting that the genes' introns were present in a common ancestor of rodents and man over 60 Myr ago. The human c-*Ha-ras-2* gene lacks introns and has not yet been biologically active on transfection (D.R.L. and E.H.C., unpublished observations). These results have suggested that c-*Ha-ras-1* may be both the more primitive and the presently active locus, while c-*Ha-ras-2* may be a pseudogene. This suggestion is further emphasized by the surprising location of c-*Ha-ras-2* on the X chromosome. The X-chromosome genes are a strictly conserved linkage group which almost never exchange or translocate to the autosomes. Ohno has suggested that the conservation of X-linked genes within mammalian orders may have resulted from the need to keep dosage-compensated genes together^{31,32}. Thus, if the prototype for c-*Ha-ras-1* were originally autosomal, the transposition to the X chromosome via an RNA intermediate was probably accompanied by rapid dysfunction to a biologically silent pseudogene. Knowledge of the relative chromosomal positions of the c-*ras* homologues in other species would clearly shed additional light on the transposition of these loci.

The c-*Ki-ras-1* gene and the c-*myb* gene are both located on chromosome 6, which is also the position of the major histocompatibility complex *HLA* in man. The participation of the *HLA* complex and the murine homologue *H-2* in leukaemia and other neoplasias is documented^{33–35}. Chromosome 6 has also been implicated as a preferred integration site for baboon endogenous retrovirus, the *BEVI* locus in man¹⁹. We cannot exclude the possibility that these different loci on chromosome 6 are physiologically as well as genetically linked. With the continued addition of many different classes of genes, including transforming genes, to the human gene map, the possibility of determining the interrelationship of various loci in neoplastic transformation becomes more realistic.

Activated *ras* genes have been found in a variety of neoplasms^{4–12,22,36}. At least two mechanisms of activation of the *Ha-ras* oncogenes have been demonstrated, the first involving the placement of *ras* under downstream promotional control of a strong retroviral LTR^{15,16}, and the second, mutation in the p21 coding sequence of c-*Ha-ras-1*^{37–39}. A potentially important variation of the above mechanisms of *onc* gene activation, is

Table 1 Chromosome and *ras* phenotype of rodent × human somatic cell hybrids

	Human chromosome																							HAS1	HAS2	K1S1	K1S2
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X				
Human	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
E36, RAG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
80H1	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	U	-	-
80H2	-	-	-	-	-	+	-	+	+	-	-	+	+	+	+	-	-	-	+	+	-	+	+	+	+	U	+
80H3	-	-	-	-	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	U	U	U
80H4	+	-	+	-	-	+	-	+	+	-	+	+	+	-	-	-	-	+	+	-	-	-	-	+	U	U	U
80H5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	U	+
80H6	-	-	-	+	-	+	-	-	-	-	-	-	+	+	+	+	-	-	-	+	+	-	-	+	U	+	U
80H7	-	+	-	-	+	-	+	-	-	-	+	+	-	+	+	-	-	+	-	+	-	-	+	+	+	U	+
80H8	-	-	+	+	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	+	-	-	U	U	U	-
80H9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	U	U	-
80H10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	U	U	-
80H11	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	U	-
80H12	-	-	-	+	-	+	-	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	+	-	+	U	-
81P1	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	+	-	U	-
81P2	-	+	+	+	-	+	+	-	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	U	+
81P4	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	+	+	U	U	U
81P5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	U	U	U
81P6	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	+	+	-	-	+	+	U	-
81P7	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	+	+	U	-
81P8	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	+	+	U	-
81P9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	+	+	U	U
81P11	-	-	-	+	-	-	-	-	+	-	-	-	+	+	+	+	-	-	-	+	-	+	+	+	-	+	U
81P12	-	-	-	+	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	+	U	-
81P13	-	-	+	-	U	U	-	+	+	-	+	U	U	+	+	-	U	U	+	+	-	-	+	+	+	U	U
81P14	-	-	-	-	-	-	-	-	+	-	-	-	+	+	+	-	-	-	-	-	-	-	+	+	+	U	-
81P15	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	+	U	+
81P16	-	-	+	+	-	-	-	+	+	+	+	-	-	+	+	-	-	+	+	+	-	+	+	+	+	U	-
81P17	-	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	-	+	+	+	+	+	+	+	+	U	-
81P18	-	-	-	+	-	-	-	-	-	-	+	-	+	-	-	+	-	+	+	+	-	-	+	+	+	U	U
70M1	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	+	U	U	-	-
70M2	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	U	U	-	-
70M3	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	U	U	+	-
70M4	+	+	+	+	-	+	+	+	+	+	+	+	-	+	-	+	+	+	+	+	+	+	+	+	U	U	+
70M5	-	-	-	-	-	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+	-	+	+	+	U	U	-
70M6	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	U	U	-	-
70M7	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	U	U	-
70M8	-	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	-	+	+	+	+	+	+	+	U	U	+
70M9	-	-	-	+	-	+	+	+	+	+	+	-	-	+	+	+	-	+	+	+	+	+	+	+	U	U	+
70M10	-	-	-	-	-	+	+	+	+	+	+	+	-	+	-	-	+	+	+	+	-	-	+	+	U	U	+
70M11	-	-	+	-	-	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	U	U	U
70M12	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	U	U	-
70M13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	U	U	-
70M14	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+	+	+	U	U	U
70M15	-	-	-	-	-	+	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	U	U	+
70M16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	U	U	+
% Discordancy with c-Ha-ras-1	50	42	38	46	48	52	50	58	46	46	0	57	52	58	54	54	52	44	38	63	38	50	42				
% Discordancy with c-Ha-ras-2	85	75	80	40	85	65	75	75	70	80	45	65	60	60	50	60	90	60	75	55	60	45	0				
% Discordancy with c-Ki-ras-1	43	43	43	35	50	0	14	36	21	29	29	43	57	14	50	36	36	29	29	57	43	21	36				
% Discordancy with c-Ki-ras-2	16	11	26	42	18	42	26	29	24	26	34	0	26	50	21	26	21	34	24	40	53	34	74				

Chromosomal and *ras* phenotype of Chinese hamster × human somatic cell hybrids. Chromosome scores indicated the consensus result of G-banding and isozyme scores. 'U' indicates not done. Thirty-six isozyme systems diagnostic for each human chromosome were run on each hybrid. The details of isozyme analysis of human hybrids have been described previously¹⁷. The isozymes and their chromosome assignments scored were: HSA 1: phosphoglucosyltransferase-1 (PGM1), 6-phosphogluconate dehydrogenase (PGD), peptidase-C (PEPC); HSA 2: isocitrate dehydrogenase (soluble) (IDH1), malate dehydrogenase (soluble) (MDH1), acid phosphatase (ACP1); HSA 3: amino acylase-1 (ACY); HSA 4: peptidase-S (PEPS), phosphoglucosyltransferase-2 (PGM2); HSA 5: hexosaminidase-B (HEXB); HSA 6: glyoxylase (GLO), phosphoglucosyltransferase-3 (PGM3), malic enzyme (ME1); HSA 7: β -glucuronidase (GUSB); HSA 8: glutathione reductase (GSR); 9 adenylate kinase (AK1); HSA 10: inorganic pyrophosphatase (PP), glutamateoxaloacetate transaminase (GOT1); HSA 11: lactate dehydrogenase-A (LDHA), acid phosphatase-2 (ACP2); HSA 12: triose phosphate isomerase (TPI), lactate dehydrogenase-B (LDHB), peptidase-B (PEPB); HSA 13: esterase-D (ESD); HSA 14: purine nucleoside phosphorylase (NP); HSA 15: mannose phosphate isomerase (MPI), hexosaminidase-A (HEXA); HSA 16: diaphorase-4 (D1A4); HSA 17: galactokinase (GALK); HSA 18: peptidase-A (PEPA); HSA 19: glucose phosphate isomerase (GPI), peptidase-D (PEPD); HSA 20: adenosine deaminase (ADA); HSA 21: superoxide dismutase-1 (SOD1); HSA 22: aconitase-2 (ACO2); HSA X: glucose 6-phosphate dehydrogenase (G6PD).

the possibility that certain human cancers may result from translocation of *onc* genes to chromosomal positions where the regulation of transcription is altered⁴⁰.

The technical assistance of Janice Martenson, Richard Bauer and Mary Eichelberger is acknowledged.

Note added in proof: Sakaguchi *et al.*⁴² have recently reported the independent genetic localization of c-Ki-ras-2 to human chromosome 12 in agreement with these findings.

Received 30 December 1982; accepted 22 February 1983.

- Cooper, G. M. *Science* **217**, 801-806 (1982).
- Klein, G. (ed.) *Advances in Viral Oncology* Vol. 1 (Raven, New York, 1982).
- Coffin, J. M. *et al. J. Virol.* **40**, 953-957 (1981).

- Shih, C., Shilo, B., Goldfarb, M. P., Dannenberg, A. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5714-5718 (1979).
- Cooper, G. M., Okenquist, S. & Silverman, L. *Nature* **284**, 418-421 (1980).
- Shih, C., Padhy, L. C., Murray, M. J. & Weinberg, R. A. *Nature* **290**, 261-264 (1981).
- Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1181-1184 (1981).
- Perucho, M. *et al. Cell* **27**, 467-476 (1981).
- Lane, M. A., Sainten, A. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5185-5189 (1981).
- Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-478 (1982).
- Santos, E., Tronick, S. R., Aaronson, S. A., Pulciano, S. & Barbacid, M. *Nature* **298**, 343-348 (1982).
- Ellis, R. W. *et al. Nature* **292**, 506-511 (1981).
- Chang, E. H., Gonda, M. A., Ellis, R. W., Scolnick, E. M. & Lowy, D. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4848-4852 (1982).
- DeFeo, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3328-3332 (1981).

16. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **297**, 479–484 (1982).
17. O'Brien, S. J., Simonson, J. M. & Eichelberger, M. in *Techniques in Somatic Cell Genetics* (ed. Shay, J. W.) 342–370 (Plenum, New York, 1982).
18. Bobro, M. & Crass, J. *Nature* **251**, 77–79 (1971).
19. Lemons, R. S., Nash, W. G., O'Brien, S. J., Benveniste, R. E. & Sherr, C. J. *Cell* **14**, 995–1005 (1978).
20. McBride, O. W. *et al. Nature* **300**, 773–774 (1982).
21. Chattopadhyay, S. K. *et al. Nature* **296**, 361–363 (1982).
22. McCoy, M. S. *et al. Nature* **302**, 79–81 (1983).
23. Shimizu, K., Goldfarb, M., Perucho, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 383–387 (1983).
24. Deisseroth, A. *et al. Cell* **12**, 205–218 (1977).
25. Deisseroth, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 1456–1460 (1978).
26. Hutton, J. J. *Biochem. Genet.* **3**, 507–515 (1969).
27. Popp, R. A. *J. Hered.* **60**, 126–133 (1969).
28. Leder, A., Swan, D., Ruddle, F., D'Eustachio, P. & Leder, P. *Nature* **293**, 196–199 (1981).
29. Popp, R. A., Lalley, P. A., Whitney, J. B. III & Anderson, W. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6362–6366 (1981).
30. Rowley, J. D. *Nature* **301**, 290–291 (1983).
31. Ohno, S. *Nature* **244**, 259–262 (1973).
32. Ohno, S. *Reproduction in Mammals* Vol. 6, 1–31 (Cambridge University Press, 1976).
33. Ryder, L. P. & Svejgaard, A. *An. Rev. Genet.* **15**, 169–187 (1981).
34. Svejgaard, A., Jersild, C., Staut-Nielsen, L. & Bodmer, W. F. *Tissue Antigens* **4**, 95–105 (1974).
35. Lilly, F. & Pincus, T. *Adv. Cancer Res.* **17**, 231 (1973).
36. Simonetta, P. *et al. Nature* **300**, 539–542 (1982).
37. Tabin, C. J. *et al. Nature* **300**, 143–149 (1982).
38. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
39. Taparowsky, E. *et al. Nature* **300**, 762–765 (1982).
40. Klein, G. *Cell* **321**, 311–315 (1983).
41. Martinville, B. de, Giacalone, J., Shih, C., Weinberg, R. A. & Franke, U. *Science* **219**, 498–500 (1983).
42. Sakaguchi, A. *et al. Science* **219**, 1081–1083 (1983).

Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant

Rik K. Wierenga & Wim G. J. Hol

Laboratory of Chemical Physics, Department of Chemistry,
University of Groningen, Nijenborgh 16, 9747 AG Groningen,
The Netherlands

Recently, it has been demonstrated that a single point mutation is responsible for the acquisition of transforming properties by the EJ and T24 human bladder carcinoma gene^{1–3}. The point mutation consists of the conversion of guanine into thymine, which results in the replacement of a glycine by a valine at position 12 of the p21 protein encoded by the EJ and T24 genes. Sequence data of retroviral analogues of the p21 protein^{1–5} also indicate the importance for a glycine residue at position 12 in normal p21. Comparison of the sequence of the 37 N-terminal residues of the normal human p21 protein with the sequence of the dinucleotide-binding $\beta\alpha\beta$ unit in a group of structurally related enzymes, suggests that these residues of p21 fold into a very similar unit which is also involved in binding a nucleotide. We present here a three-dimensional model of the p21 $\beta\alpha\beta$ unit which explains directly why glycine at position 12 cannot be replaced by another residue without altering the nucleotide-binding properties of p21.

Although the amino acid sequences of the 37 N-terminal residues of human bladder p21 protein and of several closely related oncogene products are known^{1–5}, and, furthermore, GTP binding by p21 has been established^{6–8}, there is still considerable uncertainty about the function of this important group of proteins. One approach to further our knowledge of the function of p21 is to obtain information about its three-dimensional structure. Even without performing a complete X-ray analysis, a detailed model for a protein can sometimes be constructed on the basis of the known three-dimensional structure of a protein having a related amino acid sequence^{9–11}.

As a starting point, the sequence alignment of five FAD- or NAD-binding $\beta\alpha\beta$ units is given in Table 1. This alignment, an extension of previous ones^{12–15}, is a result of a detailed study of the precise mode of binding of dinucleotides by these folding units (R.K.W., W.G.J.H. & De Maeyer, M.C.H., in preparation). The 'fingerprint' of this nucleotide-binding $\beta\alpha\beta$ unit is:

(1) the sequence Gly-X-Gly-X-X-Gly in the region joining the first β -strand and the N-terminus of the α -helix. These three glycines will be called the first, second and third essential glycine, respectively. The first and third glycine allow the polypeptide chain to make a sharp turn between the first β -strand and the α -helix. In addition, a side chain at the first glycine position would interfere with proper binding of the ribose moiety. The function of the second essential glycine has particular relevance for the rest of this paper; it provides space for a close approach of the pyrophosphate moiety to the N-terminus of the α -helix (Fig. 1A). This allows the formation of several hydrogen bonds with the protein in addition to a favourable interaction between the negative charge of the pyrophosphate group and the α -helix dipole¹⁶. Any side chain of an L-amino acid at this position would interfere with the binding of the nucleotide (Fig. 1). (2) The occurrence of several neutral, often hydrophobic, residues at the positions indicated by squares in Table 1. These residues are involved in forming the hydrophobic core of the $\beta\alpha\beta$ unit (Fig. 1A). (3) The presence of a negatively charged residue at the C-terminus of the second β -strand (symbol $\ominus$ in Table 1). This residue forms strong hydrogen bonds with the 2'-hydroxyl group of the adenine-ribose (Fig. 1A)^{12,14}. (4) The presence of a hydrophilic residue at the beginning of the first β -strand, indicated by Δ in Table 1.

Table 1 clearly shows that residues 5–33 of the p21 protein show these four characteristics almost perfectly. By using the interactive computer graphics program GUIDE¹⁷, the appropriate amino acid substitutions were made to the N-terminal $\beta\alpha\beta$ unit of the FAD-binding enzyme *p*-hydroxybenzoate hydroxylase^{18–20}, to yield the sequence of the normal human p21 protein^{1–3}. After some minor alterations of a few side chains, this gave the model of the N-terminal region of the p21 protein shown in Figs 1B and 2. Residues Leu 6, Val 8, Leu 19, Gln 22 and Val 29 form a hydrophobic core in the centre of the $\beta\alpha\beta$ unit. The only exception to the characteristics described above is the occurrence of a glutamic acid residue at position 31, where a neutral residue would be expected. This side chain, however, can form a salt bridge with Lys 16 and this exception does not disturb the $\beta\alpha\beta$ unit in any way (Fig. 1B). The binding of the ribose-pyrophosphate moiety of the nucleotide would be very similar to that in *p*-hydroxybenzoate hydroxylase (Fig. 1). A hydrogen bond is formed between a phosphate oxygen and the main-chain NH of Val 14. The C α of Gly 12 comes close to the bridging oxygen of the diphosphate. A hydrogen bond is formed between the carboxylate group of Asp 33 and the 2'-hydroxyl group of the ribose ring (Figs 1B, 2). The negatively charged residue at this position makes it unlikely that the nucleotide bound by the p21 protein is NADP as this would result in an unfavourable charge interaction between Asp 33 and the 2'-phosphate of this dinucleotide.

In view of the evidence given above, we propose that residues 5–33 of the p21 protein fold into a $\beta\alpha\beta$ unit and are essential for the binding of a nucleotide or dinucleotide, that is not NADP. Our model does not lead to a prediction with respect to the preference of the base, as a replacement of adenine by, for example, guanine (Fig. 1B) leads neither to collisions nor to additional favourable interactions. The proposed $\beta\alpha\beta$ unit would neatly explain why the presence of a glycine residue at position 12 is essential for the proper functioning of this protein: the side chain of another residue at this position would interfere with the binding of the pyrophosphate moiety of the nucleotide to the N-terminus of the α -helix. A different mode of nucleotide-binding by the p21 protein would subsequently, via several unknown steps, result in a transformed cell. It should be emphasized here that the point mutation at position 12 does not have to abolish nucleotide binding completely. It may result in an altered mode of binding of the nucleotide as has been observed for the binding of NAD analogues by liver alcohol dehydrogenase^{21,35}. In that case, steric hindrance led to a quite different position for the nicotinamide-ribose moiety of a bound NAD analogue.

Fig. 1 A, Stereo diagram of the FAD-binding $\beta\alpha\beta$ unit formed by residues 4–34 of *p*-hydroxybenzoate hydroxylase²⁰, together with the bound dinucleotide. The recently determined amino acid sequence¹⁹ has been incorporated into the original model¹⁸ and subsequently subjected to several refinement cycles. The crystallographic *R*-factor of the model is 34% for all data between 8 Å and 2.5 Å. The residues indicated by squares in Table 1 form a hydrophobic core. Residues 9, 11 and 14 are the invariant glycines of which Gly 11 is close to the pyrophosphate and other atoms of the dinucleotide. Glu 32 forms hydrogen bonds with the adenine-ribose. **B**, Stereo diagram of the proposed model for residues 5–35 of human p21 protein deduced from the structure shown in **A**. The length of this $\beta\alpha\beta$ unit is exactly the same as in *p*-hydroxybenzoate hydroxylase (Table 1) and the backbones of **A** and **B** are identical; also, the position of FAD is identical. It should be stressed that there is no evidence that the p21 protein may be involved in FAD-binding. It could be a different dinucleotide, for example NAD, or a mononucleotide such as GTP or GDP. In fact, the $\beta\alpha\beta$ unit only interacts directly with atoms from the ADP moiety of the dinucleotide. The manner in which this ADP group is bound shows that the amino acid changes necessary for constructing the p21 model from the hydroxylase $\beta\alpha\beta$ unit does not give rise to steric hindrance, but, instead, leads to several favourable interactions as discussed in the text. A few short non-bonded distances of ~2.5 Å (which should be 3.0 Å) are now being refined by energy minimization techniques; this requires only a slight movement of a few atoms.

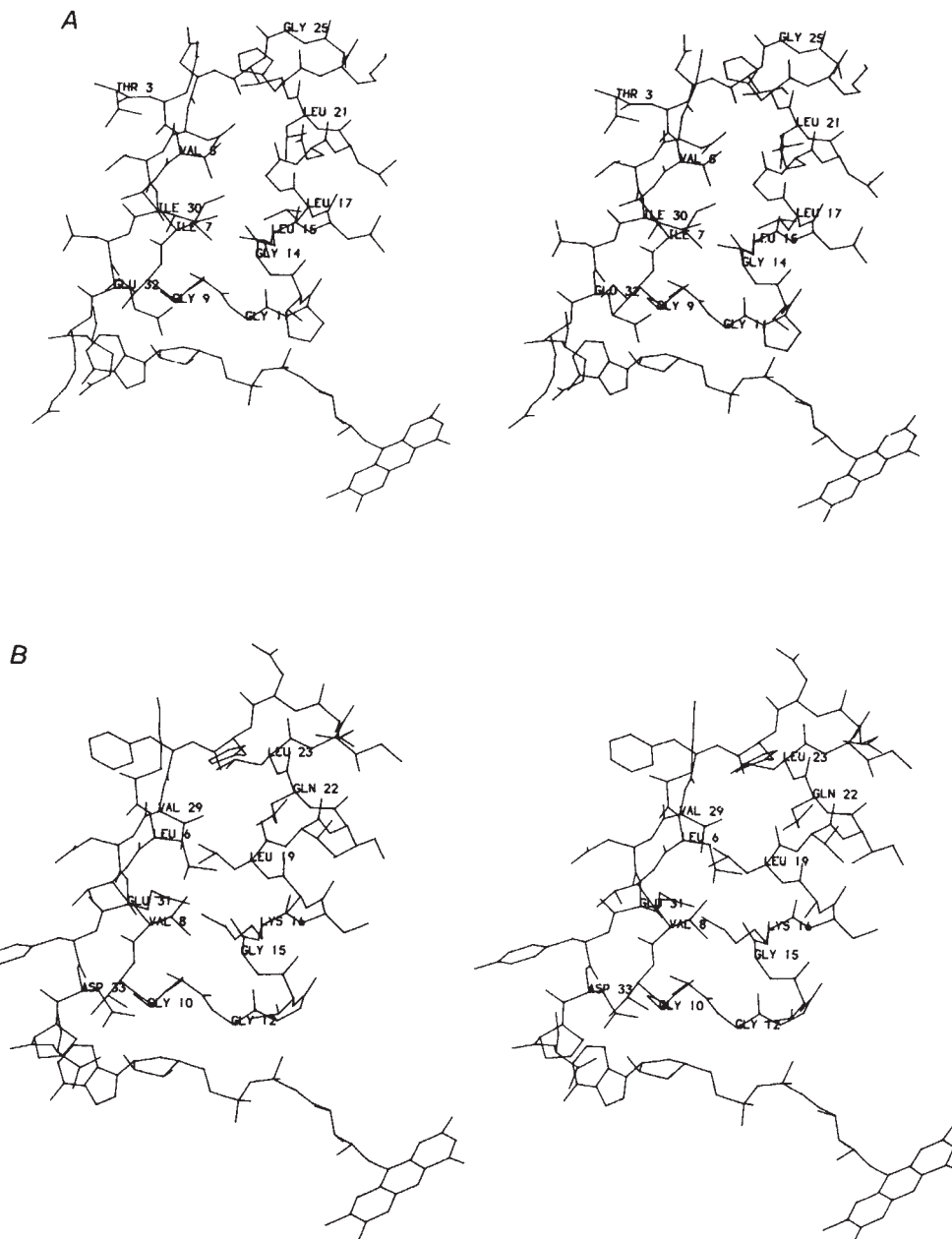


Fig. 2 Schematic drawing of the proposed binding of the ribose-diphosphate moiety of a nucleotide by the proposed N-terminal $\beta\alpha\beta$ unit of the human p21 protein (adapted from Hofsteenge *et al.*¹⁵). The non-circular symbols in the figure correspond to the 'fingerprint' symbols in Table 1. The arrows indicate the steric clash which would occur between a side chain at position 12 of the protein and several atoms of the bound nucleotide. This would lead to an altered mode of nucleotide binding and disruption of the proper functioning or control of p21.

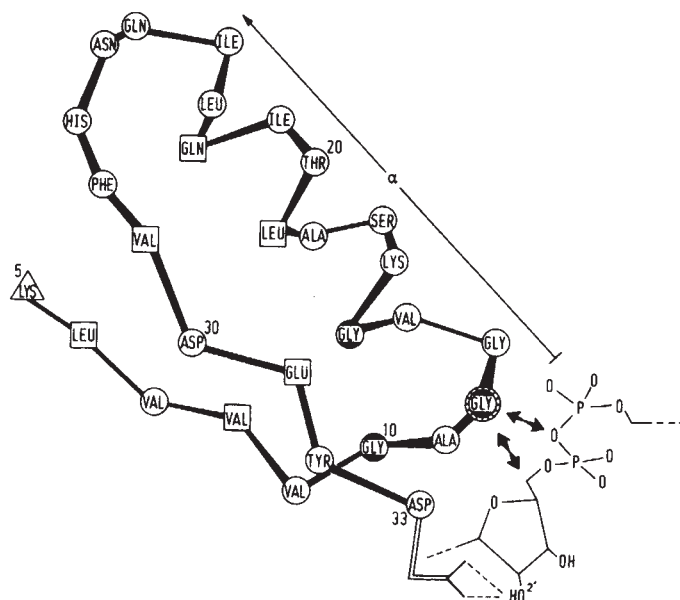


Table 1 Alignment of the amino acid sequence of normal human p21 protein with that of five dinucleotide-binding enzymes

Nucleotide		22	25	30	35	40	45	50	54	Refs							
Protein	Coenzyme																
LDH	NAD	Lys-Ile-Thr-Val-Val-Gly-Val-Gly-Ala-Val-Gly-Met-Ala-Cys-Ala-Ile-Ser-Ile-Leu-Met-Lys-Asp-Leu-Ala-Asp-Glu-Val-Ala-Leu-Val-Asp-Val-Met									28, 32						
		Δ	□	□	●	●	●	□	□	□		□	□	□	□	□	□
		194	□	□	●	200	●	●	205	□	210	□	215	□	220	□	□
ADH	NAD	Thr-Cys-Ala-Val-Phe-Gly-Leu-Gly-Gly-Val-Gly-Leu-Ser-Val-Ile-Met-Gly-Cys-Lys-Ala-Ala-Gly-Ala-Ala - Arg-Ile-Ile-Gly-Val-Asp-Ile-Asn									29, 33						
		Δ	□	□	●	●	●	□	□	□		□	□	□	□	□	□
		2	□	□	●	●	10	●	15	□	20	□	25	□	□	30	□
GPD	NAD	Lys-Ile-Gly-Ile-Asp-Gly-Phe-Gly-Arg-Ile-Gly-Arg-Leu-Val-Leu-Arg-Ala-Ala-Leu-Ser-Cys-Gly-Ala-Gln - Val-Val-Ala-Val-Asn-Asp-Pro-Phe									30, 34						
		Δ	□	□	●	●	●	□	□	□		□	□	□	□	□	□
		22	□	□	●	●	30	●	35	□	40	□	45	□	□	49	□
GR	FAD	Asp-Tyr-Leu-Val-Ile-Gly-Gly-Gly-Ser-Gly-Gly-Leu-Ala-Ser-Ala-Arg-Arg-Ala-Ala-Glu-Leu-Gly-Ala - Arg-Ala-Ala-Val-Val-Glu-Ser-His									14, 31						
		Δ	□	□	●	●	●	□	□	□		□	□	□	□	□	□
		4	□	□	●	10	●	●	15	□	20	□	25	□	□	31	□
PHBH	FAD	Gln-Val-Ala-Ile-Ile-Gly-Ala-Gly-Pro-Ser-Gly-Leu-Leu-Leu-Gly-Gln-Leu-Leu-His-Lys-Ala-Gly-Ile - Asp-Asn-Val-Ile-Leu-Glu-Arg-Gln									18, 19						
		Δ	□	□	●	●	●	□	□	□		□	□	□	□	□	□
		5	□	□	●	11	●	●	16	□	20	□	25	□	30	□	□
P21	?	Lys-Leu-Val-Val-Val-Gly-Ala-Gly-Gly-Val-Gly-Lys-Ser-Ala-Leu-Thr-Ile-Gln-Leu-Ile-Gln-Asn-His - Phe-Val-Asp-Glu-Tyr-Asn-Pro-Thr									1-3						
'Fingerprint':	Δ	□	□	●	●	●	□	□	□	□		□	□	□	□	□	□
Secondary structure	β - β - β - β - β - β	α - α - α - α - α - α - α - α - α - α - α - α - α - α - α	β - β - β - β - β - β - β														

A dash in the sequence indicates a deletion. LDH, pig lactate dehydrogenase; ADH, horse liver alcohol dehydrogenase; GPD, lobster glyceraldehyde 3-phosphate dehydrogenase; GR, human erythrocyte glutathione reductase; PHBH, *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens*. The 'fingerprint symbols' are: ●, invariant glycine; □, neutral or hydrophobic groups forming the hydrophobic core of the $\beta\alpha\beta$ unit; ⊖, the invariant negative charge involved in hydrogen bonding to a ribose hydroxyl group; Δ, invariant hydrophilic residue. The best sequence homology exists between p21 and LDH, where 10 residues out of 33 positions are identical. It is interesting that despite the lower degree of sequence homology between LDH and PHBH (three identical residues), the main-chain atoms of the $\beta\alpha\beta$ fold superimpose within 1.2 Å. This shows that the deduced fingerprint is a better guide in defining the essential features than sequence identity. As p21 follows the fingerprint for its entire length, the rather low number of four identical residues, when compared with PHBH, is of little importance and it is justifiable to build a model of p21 on the basis of PHBH coordinates, in particular as no deletions are required (Fig. 1).

Taking as a working hypothesis that GDP or GTP bind to p21 (refs 6-8) at the proposed $\beta\alpha\beta$ fold, improper binding of GTP to the oncogenic p21 might explain the autophosphorylation found for the viral protein but not for normal p21 (ref. 8). As GTP is a known allosteric effector of several proteins, one may speculate that it is also an effector of p21. Interference with proper binding of the effector would prevent the desired down-regulation of the as yet unknown function of p21. Thus, an unusually high activity would be exhibited by the modified p21 protein. This would have precisely the same dominant effect²² as a high concentration of normal p21 which has been observed in virus-induced cancer cells^{23,24}. In this manner the paradox between two apparently different causes of cancer could be resolved.

After submission of this letter, a paper by Gay and Walker²⁵ described a sequence homology between the p21 protein and mitochondrial ATP-synthase. This sequence comparison is highly interesting and shows a similar degree of homology to that which exists between p21 and LDH (Table 1). The Gay and Walker comparison can be extended to include adenylate kinase^{25,26}. The alignment of the p21 sequence with this enzyme leads to a different three-dimensional model²⁵⁻²⁷ from that shown in Fig. 1B. The analogy with adenylate kinase would predict a wide loop near Gly 12 between the first β -strand and the α -helix, where our model leads to a tight turn at this position in the $\beta\alpha\beta$ fold. In addition, the second β -strand would not form hydrogen bonds with the first β -strand which is topologically completely different from our proposed $\beta\alpha\beta$ fold. Moreover, the adenylate kinase-derived model does leave one crucial question unresolved: why is replacement of Gly 12 in the normal p21 protein by a residue having a side chain so important¹⁻³? According to the proposal of Gay and Walker²⁵, Gly 12 could be replaced by various amino acids (Arg, Ser, Pro and Glu) without a change in protein function as is shown by a comparative study of the larger family of ATP-synthase-related proteins²⁶. Central to our model is the fact, supported by structural evidence from a family of proteins, that Gly 12 cannot be changed into a residue having a side chain without p21 obtaining different properties.

We thank Mr J. van der Laan, Drs P. Prick, M. de Maeyer and J. Drenth for assistance and stimulating discussions; and the editor and referees for important suggestions. This research was supported, in part, by the Dutch Foundation for Chemical Research (SON) with financial aid from the Dutch Organisation for the Advancement of Pure Research (ZWO).

Received 12 January; accepted 8 March 1983.

1. Tabin, C. J. *et al.* *Nature* **300**, 143-149 (1982).
2. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149-152 (1982).
3. Taparowsky, E. *et al.* *Nature* **300**, 762-765 (1982).
4. Dhar, R. *et al.* *Science* **217**, 934-936 (1982).
5. Tsuchida, N., Ryder, T. & Ohtsubo, E. *Science* **217**, 937-939 (1982).
6. Scolnick, E. M., Papageorge, A. G. & Shih, T. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5355-5359 (1979).
7. Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O. & Scolnick, E. M. *Nature* **287**, 686-691 (1980).
8. Papageorge, A., Lowy, D. & Scolnick, E. M. *J. Virol.* **44**, 509-519 (1982).
9. Bedarkar, S., Turnell, W. G., Blundell, T. L. & Schwabe, C. *Nature* **270**, 449-451 (1977).
10. Isaacs, N. *et al.* *Nature* **271**, 278-281 (1978).
11. Blundell, T. L., Bedarkar, S., Rinderknecht, E. & Humbel, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 180-184 (1978).
12. Ohlsson, I., Nordström, B. & Brändén, C. I. *J. molec. Biol.* **89**, 339-354 (1974).
13. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250**, 194-199 (1974).
14. Schulz, G. E., Schirmer, R. H. & Pai, E. F. *J. molec. Biol.* **160**, 287-308 (1982).
15. Hofsteenge, J. *et al.* *Eur. J. Biochem.* **113**, 141-150 (1980).
16. Hol, W. G. J., Van Duijnen, P. Th. & Berendsen, H. J. C. *Nature* **273**, 443-446 (1978).
17. Brandenburg, N. P., Dempsey, S., Dijkstra, B. W., Lijk, L. J. & Hol, W. G. J. *J. appl. Crystallogr.* **14**, 274-279 (1981).
18. Wierenga, R. K., De Jong, R. J., Kalk, K. H., Hol, W. G. J. & Drenth, J. *J. molec. Biol.* **131**, 55-73 (1979).
19. Weijer, W. J., Hofsteenge, J., Vereijken, J. M., Jekel, P. A. & Beinbema, J. *J. Biochim. Biophys. Acta* **704**, 385-388 (1982).
20. Müller, F. & Van Berkel, W. J. H. *Eur. J. Biochem.* **128**, 21-27 (1982).
21. Samama, J. P., Zeppezauer, E., Biellmann, J. F. & Brändén, C. I. *Eur. J. Biochem.* **81**, 403-409 (1977).
22. Cooper, G. M. *Science* **218**, 801-806 (1982).
23. DeFeo, D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 3328-3332 (1981).
24. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **297**, 479-483 (1982).
25. Gay, N. J. & Walker, J. E. *Nature* **301**, 262-264 (1983).
26. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. *EMBO J.* **1**, 945-951 (1982).
27. Pai, E. F., Sachsenheimer, W., Schirmer, R. H. & Schulz, G. E. *J. molec. Biol.* **114**, 37-45 (1977).
28. White, J. L. *et al.* *J. molec. Biol.* **102**, 759-779 (1976).
29. Eklund, H. *et al.* *J. molec. Biol.* **146**, 561-587 (1981).
30. Moras, D. *et al.* *J. biol. Chem.* **250**, 9137-9162 (1975).
31. Krauth-Sieghel, R. L. *et al.* *Eur. J. Biochem.* **121**, 259-267 (1982).
32. Taylor, S. S. *J. biol. Chem.* **252**, 1799-1806 (1977).
33. Jorvall, H. *Eur. J. Biochem.* **16**, 25-40 (1970).
34. Davidson, B. E., Sajó, M., Noller, H. F. & Harris, J. I. *Nature* **216**, 1181-1185 (1967).
35. Samama, J. P., Wrixon, A. D. & Biellmann, J. F. *Eur. J. Biochem.* **118**, 479-486 (1981).